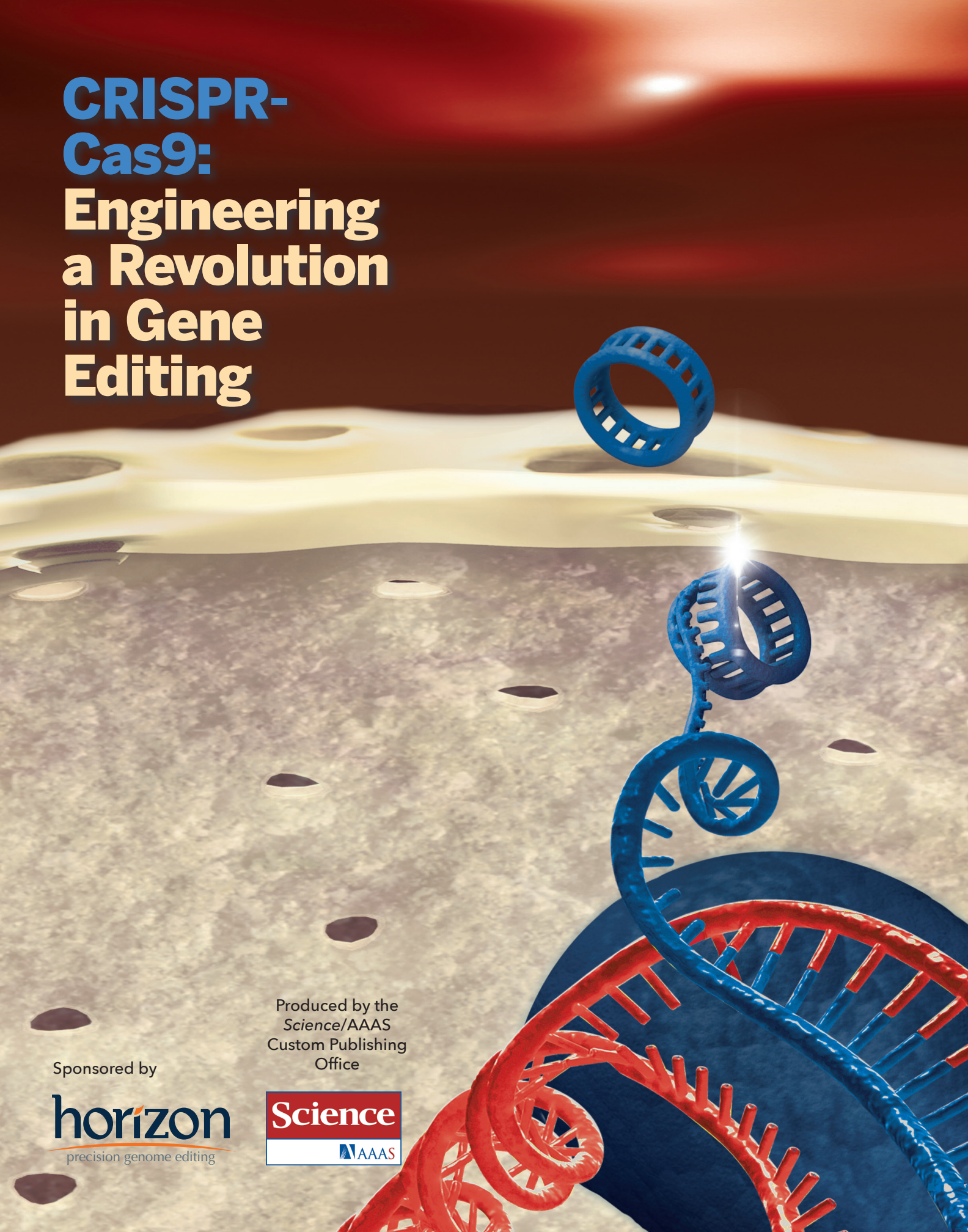
CRISPR-Cas9: Engineering a Revolution in Gene Editing



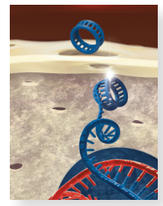
Produced by the *Science*/AAAS Custom Publishing Office

Sponsored by

horizon
precision genome editing

Science
AAAS

CRISPR-Cas9: Engineering a Revolution in Gene Editing



Cover image provided courtesy of Horizon Discovery. This image has been digitally altered from the original version for publication purposes.

Editors: Tianna Hicklin, Ph.D. and Sean Sanders, Ph.D.; Designer: Amy Hardcastle; Copyeditor: Yuse Lajiminmuhip

This booklet was produced by the *Science*/AAAS Custom Publishing Office and supported by Horizon Discovery.

The start of a new genomic era

Science has been at the forefront of publishing some of the groundbreaking work as scientists have begun to unravel the CRISPR-Cas9 system.

Humans are an exceptionally adaptive species—a characteristic that has enabled us to flourish all over the planet. We have adapted genetically and epigenetically to many different climates and habitats, and these adaptive mutations have been passed down to subsequent generations. However, these are not the only mechanisms at play; we also have thrived by modifying our surroundings, passing this information on to the next generation so that it can be built upon, refined, and improved. Now, technological advances in genomic engineering hold the potential to give us the key to not only modifying our external environment, but to also engineering genetic adaptations for ourselves as well as other species.

Starting with the discovery of mysterious palindromic, repeated DNA sequences in *E. Coli* in 1987, scientists began investigating the function of this seemingly odd phenomenon. Out of this natural curiosity grew an entirely new way to modify DNA: CRISPR-Cas9. The system evolved as a self-defense mechanism for bacteria—essentially, a way to self-vaccinate against invading viruses and plasmids (*Science* 23 March 2007, p. 1709, scim.ag/1oSz2rE). This "adaptive immunity," has enabled bacteria and archaea to continually arm and rearm themselves against invaders. However, now that scientists have taken the reins, the CRISPR-Cas9 system has been retooled into a more globally viable technology, whereby the genetic code of virtually any species can be modified, and for more than simply self-protection.

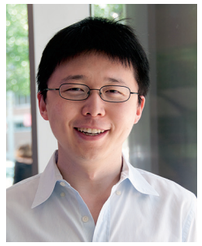
Named as a runner up for *Science* Magazine's "Breakthroughs of the Year" in 2013 (*Science* 20 December 2013, p. 1434*;* scim.ag/1uykTTC), the CRISPR-Cas9 system is revolutionizing genomic engineering and equipping scientists with the ability to precisely modify the DNA of essentially any organism. This gene editing could potentially confer genetic advantages that previously took large amounts of evolutionary time (and perhaps a bit of luck), taxing genetic breeding strategies, or bulkier and more complex genomic editing tools to acquire. Speculations about what's on the horizon seem to be limitless at this point.

Just how powerful is this technique? The ability for precision genome engineering comes with the potential to enhance food production, medicinal discoveries, and energy solutions, to name a few. Studies over the past several years have shown promise for altering crop resistance to infection and disease, advancing drug discovery, modifying fuel/energy sources, and even elucidating the multiple genetic contributions underlying human diseases—from heart disease to mental illnesses.

Science has been at the forefront of publishing some of the groundbreaking work as scientists have begun to unravel the CRISPR-Cas9 system and invent novel ways to use this tool. Two such papers were the seminal work of Feng Zhang and George Church, which were simultaneously published and the first studies to show that the CRISPR-Cas9 system can alter mammalian genomes, including humans (see pages 20, 24). Since that time, a "CRISPR Craze" has begun (see page 17), whereby scientists have begun exploring more and more ingenious ways to use this technology. Researchers have already progressed from studies on individual DNA alterations into studies using CRISPR-Cas9 as an "efficient, large-scale, loss-of-function screening method in mammalian cells," in lieu of using RNAi screens (see pages 9, 13).

Of course, it is hoped that these types of studies are just the tip of the iceberg for this new era of precision genetic engineering. In this booklet, we invite the reader to explore a selection of articles that highlight both the history of this technique and how it has grown into one of the most powerful and precise genomic engineering tools to date.

Tianna Hicklin, Ph.D.
Editor, *Science*/AAAS Custom Publishing Office



CRISPR: The democratization of gene editing

The elegance and simplicity of Cas9 have sparked the imagination of scientists across many scientific disciplines.

Recent development of genome editing technologies based on the RNA-guided CRISPR-associated endonuclease Cas9 has generated enormous excitement across many fields, including biological research, biotechnology, and medicine. For the first time, researchers have gained the ability to achieve targeted genomic modifications with efficiency and ease. This is particularly true when combined with the rapidly increasing amount of information available from genomic sequencing efforts available as well as innovative nucleic acid synthesis and delivery systems.

Unlike previous generations of genome editing tools based on zinc finger and transcription activator-like effector proteins, which achieve sequence recognition via protein-DNA interactions, Cas9 can be targeted to specific genomic loci with a guide RNA (gRNA). Once the Cas9-gRNA complex finds the DNA target via Watson-Crick base pairing, Cas9 introduces a double-strand DNA break at the target site, which in turn catalyzes targeted genome editing via non-homologous end joining or homology directed repair. The ability to use nucleic-acid hybridization rules to reprogram Cas9 specificity significantly simplifies genome editing applications particularly given that gRNAs are easily synthesized and introduced into cells to facilitate targeted genome modifications.

Despite being a nascent technology, Cas9 has been successfully used to generate an increasing number of cellular and animal models for a variety of basic research as well as biotechnology applications. For example, Cas9 can facilitate the generation of isogenic cell lines to identify causal genetic variations. In addition, Cas9 has already been broadly applied in many species to generate transgenic models, including mouse, rat, zebrafish, fruit flies, *C. elegans*, primates, and a variety of plant species. For each species, gRNAs can be easily designed based on reference genome sequences to target virtually any locus of choice. Moreover, direct application of Cas9 in embryos can significantly accelerate transgenic manipulation of whole organisms, including many previously intractable species.

In addition to facilitating the editing of individual genes, Cas9 can also be used for high throughput genetic screening applications where many genes are perturbed in a multiplexed fashion. Large libraries of gRNAs capable of targeting wide ranges or subsets of genes in the human genome can be easily synthesized using oligo array synthesis and used to generate genome-scale gene knockout libraries for functional screening. For example, these Cas9 gRNA libraries have already been successfully used in a number of studies to identify drug resistance genes. In the future, Cas9 gRNA libraries may also be used to facilitate high throughput investigation of non-coding sequences by disrupting regulatory regions in the genome.

Cas9 can also be converted into a catalytically inactive RNA-guided DNA binding protein, which can be used to recruit transcription effector domains to specific genomic loci to modulate transcription state. Alternatively, catalytically inactive Cas9 can also be linked to fluorescent proteins to enable direct visualization of genomic loci in living cells.

The pace of Cas9 development is only accelerating. Researchers are rapidly enhancing the functionalities of Cas9, making it more specific, efficient, and easier to use in a variety of biological contexts. At the same time, the elegance and simplicity of Cas9 have sparked the imagination of scientists across many scientific disciplines, with many of them already using the technology to uncover fundamental biological processes and develop innovative therapeutic strategies for treating intractable diseases. With all of these developments, the road ahead is undoubtedly full of exciting possibilities.

Dr. Feng Zhang, Ph.D., is currently serving as the W.M. Keck Career Development Professor with a joint appointment in the Biological Engineering and Brain and Cognitive Sciences Departments at the Massachusetts Institute of Technology (MIT), is one of the 11 core members of the Broad Institute of MIT and Harvard and is an Investigator at the McGovern Institute for Brain Research.

EDITORIAL

The power and possibilities of genome engineering

By Jeffrey M. Perkel

Ever since scientists understood that DNA carried heritable information, they have desired to bend this code to their will. Manipulating the fundamental code for life would mean the ability to correct defects and permanently cure genetic disorders. But the tools that have been available up until now, although workable, are crude at best—like trying to perform surgery while wearing mittens.

These days, though, the metaphorical mittens are off. Over the past decade, researchers have devised a succession of strategies that can handle just about any genetic rewrites they can imagine. These "genome-editing" technologies aren't perfect, and they are typically limited to one or just a few changes at a time. But they do fill a crucial hole in biologists' toolboxes. In so doing, they are redefining the genetic frontier, and redefining it for the better.

Genome editing technologies exist in a handful of basic forms, one of which is based on homologous recombination. In general, homologous recombination rates are too low in mammalian cells to make this approach practical (a key exception being mouse embryonic stem cells, which is why mouse transgenic technology has been so successful). By inserting the repair template into a single-stranded recombinant adeno-associated viral (rAAV) backbone however, efficiency improves considerably to the point that the approach is workable to generate precise genomic modifications.

Most genome-editing strategies, however, rely on the concept of a customized DNA-cutting endonuclease. Several variants of this approach have been developed, but in all these cases, the idea is the same: To generate a double-stranded DNA break at a specified location in the genome of a live cell.

In responding to that break, the cell may successfully stitch the two ends together (i.e., no change to the sequence occurs); it may inadvertently disable the gene by non-homologous end joining (NHEJ)-based frame-shifts; or it may repair the gene via homology-directed repair (HDR), a process that researchers can game by supplying their own templates to repair a point mutation or insert a missing gene.

That template molecule typically is supplied either in the form of a short oligonucleotide or double-stranded DNA plasmid. A newer strategy uses a rAAV vector instead, which combines the enhanced nuclear uptake of a single-stranded molecule with the greater donor length of plasmids.

In its original implementation, genome editing was accomplished using meganucleases, a class of natural nucleases that targets relatively long DNA recognition sequences. But these enzymes proved difficult to tailor, requiring either mutagenesis or sophisticated protein engineering.

Zinc finger nucleases (ZFNs) introduced the concept of programmability to genome editing. ZFNs fuse the DNA-binding domains of zinc finger transcription factors to a generic DNA endonuclease to induce a double-stranded DNA break at a defined position. Programming is achieved by selecting the identity and order of fingers used.

In a zinc finger transcription factor, each "finger" recognizes a specific three- or four-base sequence. By simply stringing together an appropriate array of fingers, researchers can in theory target any sequence they desire, and researchers have been successful at doing precisely that. Sangamo BioSciences, for instance, has demonstrated it can safely use ZFN technology in human patients. In a study published earlier this year in the *New England Journal of Medicine*, the first ever to document genome-engineering in the clinic, researchers knocked out the gene for the HIV co-receptor, CCR5, in T cells from HIV+ individuals, and then safely returned those cells to the patients, raising T cell counts overall and CD4 counts in particular (*1*).

Still, relatively few researchers have hitched their wagon to ZFN technology, as making high-quality ZFNs takes considerable skill. One key problem: Individual fingers don't always function as expected in the context of an intact ZFN. Optimization—and therefore additional time and money—is often required.

Another class of artificial enzymes, the transcription activator-like (TAL) effector nucleases (TALENs), deliver similar benefits to ZFNs yet generally are easier to use. Like ZFNs, TALENs are modular transcription factors, with each so-called TAL module specifying a single base in the recognition sequence. As TAL modules function more or less like reusable building blocks, TALENs are far simpler to design than ZFNs and can be assembled from kits relatively quickly. Yet their size and repetitive nature can pose challenges for the uninitiated.

The newest genome-editing strategy is the so-called CRISPR-Cas9 system. CRISPR-Cas9 is not an artificial construction; bacteria use it as a primitive form of adaptive anti-viral immunity. When infected with a pathogen, bacteria endowed with this system retain a signature of the infecting agent in their chromosomal DNA. The cell then transcribes those sequences—called CRISPRs, or "clustered regularly interspaced short palindromic repeats"—and processes them into short RNAs called crRNAs. With a second transcript called a tracrRNA, the crRNA guides the Cas9 nuclease to its target—normally, an invading viral nucleic acid.

In a seminal 2012 article in *Science*, Jennifer Doudna and Emmanuelle Charpentier worked out the mechanism of this process, and showed that it could be both reprogrammed and simplified into a two-component system in vitro (*2*). Subsequent studies showed that the system could be harnessed to hit genomic sequences in bacteria, fruit flies, nematodes, and human cells.

CRISPR-Cas9 offers several key benefits over competing endonuclease technologies. First, while meganucleases, ZFNs, and TALENs can be thought of as bespoke single-function machines, Cas9 is basically a programmable enzyme. All that is required is a construct expressing the generic Cas9 nuclease and a set of instructions in the form of a "single-guide RNA" (sgRNA) complementary to the desired target. The system is simple and inexpensive to implement and thus more attractive to researchers who might have been skittish of ZFNs and TALENs.

CRISPR-Cas9 is also relatively efficient, editing target sequences at surprisingly high rates.

The second benefit is multiplexing. Since Cas9 is guided by its sgRNA, researchers can program it with multiple guide RNAs simultaneously. Feng Zhang and George Church, writing independently in *Science*, have both demonstrated the ability to target two sites simultaneously (*3, 4*), and Rudolf Jaenisch has targeted five (*5*).

CRISPR-Cas9 is also relatively efficient, editing target sequences at surprisingly high rates. As a general rule, HDR-mediated genome editing typically occurs at much lower frequencies than NHEJ, requiring as it does a second piece of DNA containing the repair template. Similarly, it is easier to hit one allele than two. Yet Jaenisch found that 20 of 96 mouse embryonic stem cell clones tested using three sgRNAs simultaneously contained NHEJ-induced mutations at all six alleles of those three genes, a 20% success rate (*5*). Church observed HDR-mediated repair rates in human cells of 3% and 8% using two separate sgRNAs, compared to about 0.5% with a TALEN directed at the same location (*4*).

That's not to say CRISPR-Cas9 is perfect. Multiple studies have documented off-site targeting when using the system, for instance, at least in its original incarnation—something that could significantly limit potential clinical applications. Researchers have developed strategies to boost targeting specificity,

including both paired nickase and FokI-fusion approaches that require two closely spaced sgRNA binding events for cleavage, but whether they will be sufficient to make CRISPR-Cas9 clinically useful remains an open question.

One particularly promising application of genome editing marries its power with induced pluripotent stem cell (iPS) technology.

Still, there's little doubt the system will be useful in the research lab. Indeed, researchers are beginning to explore its application to more sophisticated genomic operations. Echoing work previously done with ZFNs and TALENs, for instance, they have made surgical changes in gene expression (as opposed to sequence) by coupling the catalytically inactive Cas9 enzyme to transcriptional regulatory domains and illuminated chromosome structure by coupling it to fluorescent proteins. George Church has experimented with orthologous Cas9 proteins to expand the technology's reach even further, for instance by allowing combined genome editing and transcriptional regulation (*6*).

Indeed, researchers today are extending the CRISPR-Cas9 system on multiple fronts. Teams led independently by Feng Zhang and Eric Lander, writing in *Science*, recently demonstrated the utility of the system for large-scale screening, much as researchers did a decade ago using RNA interference (*7, 8*). Another team used the system to build a set of biological circuits for synthetic biology applications (*9*). And still other researchers are exploring the clinical applications of the technology, through startup firms such as Editas Medicine and CRISPR Therapeutics.

One recent study, led by Daniel Anderson at Massachusetts Institute of Technology (MIT), illustrates the potential power of CRISPR-Cas9 in the clinic. Anderson's team co-delivered an sgRNA, Cas9, and a 199-base single-stranded donor template molecule into the tail vein of a mouse model of a genetic disease called hereditary tyrosinemia type I, caused by a single point mutation in the gene for fumarylacetoacetate hydrolase. Though such a strategy cannot directly be implemented in humans, the treatment did repair the mutation in one in every 250 liver cells in the treated animals, minimizing the liver damage and weight loss typically seen in this model (*10*).

One particularly promising application of genome editing marries its power with induced pluripotent stem cell (iPS) technology. iPS cells, in which researchers turn a somatic cell into an embryonic-like pluripotent stem cell using four transcription factors and/or small molecules, are opening the door not only to new studies of disease etiology, but corrective cellular therapies as well, as evidenced by a new iPS cell-based clinical study launched recently in Japan.

That study hopes to differentiate patient-specific iPS cells into retinal pigment epithelial cells for the treatment of age-related macular degeneration. But combined with genome-editing technology, iPS cells can do much, much more—reconstituting a sickle-cell anemia patient's bone marrow with hematopoietic stem cells in which the mutant hemoglobin gene has been repaired, for instance. Sangamo BioSciences demonstrated recently that it could use ZFNs to drive HDR-based repair of the IL2R-gamma gene in human hematopoietic stem cells, but that work was not done in a clinical setting (*11*).

Where ZFNs, or any genome editing technology for that matter, will go from here is an open question. But given the pace of research and the steady stream of promising findings so far, it's a good bet that it's going to be an exciting ride.

References

1. P. Tebas *et al.*, *NEJM* **370**, 901 (2014).
2. M. Jinek *et al.*, *Science* **337**, 816 (2012).
3. L. Cong *et al.*, *Science* **339**, 819 (2013).
4. P. Mali *et al.*, *Science* **339**, 823 (2013).
5. H. Wang *et al.*, *Cell* **153**, 910 (2013).
6. K.M. Esvelt *et al.*, *Nat. Methods* **10**, 1116 (2013).
7. O. Shalem *et al.*, *Science* **343**, 84 (2014).
8. T. Wang *et al.*, *Science* **343**, 80 (2014).
9. S. Kiani *et al.*, *Nat. Methods* **11**, 723 (2014).
10. H. Yin *et al.*, *Nat. Biotechnol.* **32**, 551 (2014).
11. P. Genovese *et al., Nature* **510**, 235 (2014); 10.1038/nature13420.

Jeffrey M. Perkel is a freelance science writer based in Pocatello, Idaho.

RESEARCH

Genetic screens in human cells using the CRISPR-Cas9 system

Tim Wang,[1,2,3,4] Jenny J. Wei,[1,2] David M. Sabatini,[1,2,3,4,5*†] Eric S. Lander[1,3,6*†]

The bacterial clustered regularly interspaced short palindromic repeats (CRISPR)–Cas9 system for genome editing has greatly expanded the toolbox for mammalian genetics, enabling the rapid generation of isogenic cell lines and mice with modified alleles. Here, we describe a pooled, loss-of-function genetic screening approach suitable for both positive and negative selection that uses a genome-scale lentiviral single-guide RNA (sgRNA) library. sgRNA expression cassettes were stably integrated into the genome, which enabled a complex mutant pool to be tracked by massively parallel sequencing. We used a library containing 73,000 sgRNAs to generate knockout collections and performed screens in two human cell lines. A screen for resistance to the nucleotide analog 6-thioguanine identified all expected members of the DNA mismatch repair pathway, whereas another for the DNA topoisomerase II (*TOP2A*) poison etoposide identified *TOP2A*, as expected, and also cyclin-dependent kinase 6, *CDK6*. A negative selection screen for essential genes identified numerous gene sets corresponding to fundamental processes. Last, we show that sgRNA efficiency is associated with specific sequence motifs, enabling the prediction of more effective sgRNAs. Collectively, these results establish Cas9/sgRNA screens as a powerful tool for systematic genetic analysis in mammalian cells.

A critical need in biology is the ability to efficiently identify the set of genes underlying a cellular process. In microorganisms, powerful methods allow systematic loss-of-function genetic screening (*1*, *2*). In mammalian cells, however, current screening methods fall short—primarily because of the difficulty of inactivating both copies of a gene in a diploid mammalian cell. Insertional mutagenesis screens in cell lines that are near-haploid or carry *Blm* mutations, which cause frequent somatic crossing-over, have proven powerful but are not applicable to most cell lines and suffer from integration biases of the insertion vectors (*3*, *4*). The primary solution has been to target mRNAs with RNA interference (RNAi) (*5–9*). However, this approach is also imperfect because it only partially suppresses target gene levels and can have off-target effects on other mRNAs, resulting in false negative and false positive results (*10–12*). Thus, there remains an unmet need for an efficient, large-scale, loss-of-function screening method in mammalian cells.

Recently, the clustered regularly interspaced short palindromic repeats (CRISPR) pathway, which functions as an adaptive immune system in bacteria (*13*), has been co-opted to engineer mammalian genomes in an efficient manner (*14–16*). In this two-component system, a single-guide RNA (sgRNA) directs the Cas9 nuclease to cause double-stranded cleavage of matching target DNA sequences (*17*). In contrast to previous genome-editing techniques, such as zinc-finger nucleases and transcription activator-like effector nucleases (TALENs), the target specificity of CRISPR-Cas9 is dictated by a 20–base pair (bp) sequence at the 5′ end of the sgRNA, allowing for much greater ease of construction of knockout reagents. Mutant cells lines and mice bearing multiple modified alleles can be generated with this technology (*18*, *19*).

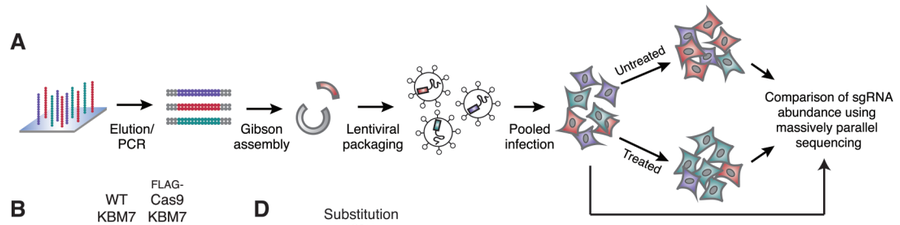
We set out to explore the feasibility of using the CRISPR-Cas9 system to perform large-scale, loss-of-function screens in mammalian cells. The idea was to use a pool of sgRNA-expressing lentivirus to generate a library of knockout cells that could be screened under both positive and negative selection. Each sgRNA would serve as a distinct DNA barcode that can be used to count the number of cells carrying it by using high-throughput sequencing (Fig. 1A). Pooled screening requires that single-copy sgRNA integrants are sufficient to induce efficient cleavage of both copies of a targeted locus. This contrasts with the high expression of sgRNAs achieved through transfection that is typically used to engineer a specific genomic change by using the CRISPR-Cas9 system.

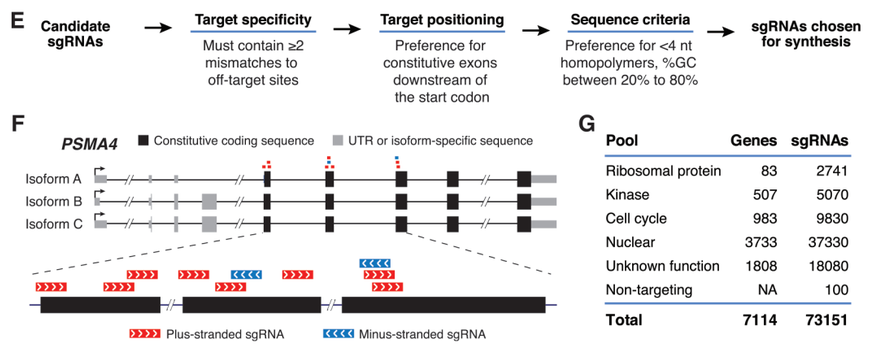
We first tested the concept in the near-haploid, human KBM7 CML cell line by creating a clonal derivative expressing the Cas9 nuclease (with a FLAG-tag at its N terminus) under a doxycycline-inducible promoter (Fig. 1B). Transduction of these cells at low multiplicity of infection (MOI) with a lentivirus expressing a sgRNA targeting the endogenous AAVS1 locus revealed substantial cleavage at the AAVS1 locus 48 hours after infection (Fig. 1C). Moreover, because the sgRNA was stably expressed, genomic cleavage continued to increase over the course of the experiment. Deep sequencing of the locus revealed that repair of Cas9-induced double-strand breaks resulted in small deletions (<20 bp) in the target sequence, with tiny insertions or substitutions (<3 bp) occurring at a lower frequency (Fig. 1D). The vast majority of the lesions, occurring in a protein-coding region, would be predicted to give rise to a nonfunctional protein product, indicating that CRISPR-Cas9 is an efficient means of generating loss-of-function alleles.

We also analyzed off-target activity of CRISPR-Cas9. Although the specificity of CRISPR-Cas9 has been extensively characterized in transfection-based settings (*20–22*), we wanted to examine its off-target behavior in our system, in which Cas9 and a sgRNA targeting AAVS1 (sgAAVS1) were stably expressed for 2 weeks. We compared the level of cleavage observed at the target locus (97%) with levels at 13 potential off-target cleavage sites in the genome (defined as sites differing by up to 3 bp from sgAAVS1) (fig. S1A). Minimal cleavage (<2.5%) was observed at all sites, with one exception, which was the only site that had perfect complementarity in the "seed" region (terminal 8 bp) (fig. S1B). On average, sgRNAs have ~2.2 such sites in the genome, almost always (as in this case) occurring in noncoding DNA and thus less likely to affect gene function (supplementary text S1).

To test the ability to simultaneously screen tens of thousands of sgRNAs, we designed a sgRNA library with 73,151 members, consisting of multiple sgRNAs targeting 7114 genes and 100 nontargeting controls (Fig. 1E, table S1, and supplementary materials, materials and methods). sgRNAs were designed against constitutive coding exons near the beginning of each gene and filtered for potential off-target effects based on sequence similarity to the rest of the human genome (Fig. 1, F and G). The library included 10 sgRNAs for each of 7031 genes and all possible sgRNAs for each of the 83 genes encoding ribosomal proteins (Fig. 1H). To assess the effective representation of our microarray synthesized library, we sequenced sgRNA barcodes from KBM7 cells 24 hours after infection with the entire lentiviral pool and were able to detect the overwhelming majority (>99%) of our sgRNAs, with high uniformity across constructs (only a sixfold increase in abundance

[1]Department of Biology, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA.
[2]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA.
[3]Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA.
[4]David H. Koch Institute for Integrative Cancer Research at MIT, Cambridge, MA 02139, USA.
[5]Howard Hughes Medical Institute, Department of Biology, MIT, Cambridge, MA 02139, USA.
[6]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA.
† These authors contributed equally to this work.
***Corresponding author. E-mail: sabatini@wi.mit.edu (D.M.S.); lander@broadinstitute.org (E.S.L.)**





G

Pool	Genes	sgRNAs
Ribosomal protein	83	2741
Kinase	507	5070
Cell cycle	983	9830
Nuclear	3733	37330
Unknown function	1808	18080
Non-targeting	NA	100
Total	**7114**	**73151**

Fig. 1. A pooled approach for genetic screening in mammalian cells by using a lentiviral CRISPR-Cas9 system. (**A**) Outline of sgRNA library construction and genetic screening strategy. (**B**) Immunoblot analysis of WT KBM7 cells and KBM7 cells transduced with a doxycycline-inducible FLAG-Cas9 construct upon doxycycline induction. S6K1 was used as a loading control. (**C**) Sufficiency of single-copy sgRNAs to induce genomic cleavage. Cas9-expressing KBM7 cells were transduced with AAVS1-targeting sgRNA lentivirus at low MOI. The SURVEYOR mutation detection assay was performed on cells at the indicated days post-infection (dpi). Briefly, mutations resulting from cleavage of the AAVS1 locus were detected through polymerase chain reaction (PCR) amplification of a 500-bp amplicon flanking the target sequence, re-annealing of the PCR product, and selective digestion of mismatched heteroduplex fragments. (**D**) Characterization of mutations induced by CRISPR-Cas9 as analyzed with high-throughput sequencing. (**E**) sgRNA library design pipeline. (**F**) Example of sgRNAs designed for *PSMA4*. sgRNAs targeting constitutive exonic coding sequences nearest to the start codon were chosen for construction. (**G**) Composition of genome-scale sgRNA library.

between the 10th and 90th percentiles) (fig. S2A).

As an initial test of our approach, we screened the library for genes that function in DNA mismatch repair (MMR). In the presence of the nucleotide analog 6-thioguianine (6-TG), MMR-proficient cells are unable to repair 6-TG–induced lesions and arrest at the G2-M cell-cycle checkpoint, whereas MMR-defective cells do not recognize the lesions and continue to divide (*23*). We infected Cas9-KBM7 cells with the entire sgRNA library, cultured the cells in a concentration of 6-TG that is lethal to wild-type (WT) KBM7 cells, and sequenced the sgRNA barcodes in the final population. sgRNAs targeting the genes encoding the four components of the MMR pathway (*MSH2*, *MSH6*, *MLH1*, and *PMS2*) (*24*) were dramatically enriched in the 6-TG–treated cells. At least four independent sgRNAs for each gene showed very strong enrichment, and barcodes corresponding to these genes made up >30% of all barcodes (Fig. 2, A and B). Each of the 20 most abundant sgRNAs targeted one of these four genes. The fact that few of the other 73,000 sgRNAs scored highly in this assay suggests a low frequency of off-target effects.

We next addressed the challenge of loss-of-function screening in diploid cells, which require biallelic inactivation of a target gene. We therefore generated an inducible Cas9 derivative of the HL60 pseudo-diploid human leukemic cell line. In both HL60 and KBM7 cells, we screened for genes whose loss conferred resistance to etoposide, a chemotherapeutic agent that poisons DNA topoisomerase IIA (*TOP2A*). To identify hit genes, we calculated the difference in abundance between the treated and untreated populations for each sgRNA, calculated a score for each gene using a Kolmogorov-Smirnov test to compare the sgRNAs targeting the gene against the nontargeting control sgRNAs, and corrected for multiple hypothesis testing (Fig. 2, C to E, and table S2). Identical genes were detected in both screens, with significance levels exceeding all other genes by more than 100-fold. As expected, loss of *TOP2A* itself conferred strong protection to etoposide (*25*). The screen also revealed a role for *CDK6*, a G_1 cyclin-dependent kinase, in mediating etoposide-induced cytotoxicity. Every one of the 20 sgRNAs in the library targeting *TOP2A* or *CDK6* was strongly enriched (>90th percentile) in both screens, indicating that the effective coverage of our libraries is very high. We generated isogenic HL60 cell lines with individual sgRNAs against *TOP2A* and *CDK6* and, consistent with the screen results, these lines were much more resistant to etoposide than parental or sgAAVS1-modified HL60 cells (Fig. 2, F and G). Thus, our Cas9/sgRNA system enables large-scale positive selection loss-of-function screens.

To identify genes required for cellular proliferation, we screened for genes whose loss conferred a selective disadvantage on cells. Such a screen requires accurate identification of sgRNAs that are depleted from the final cell population. A sgRNA will show depletion only if cleavage of the target gene occurs in the majority of cells carrying the construct.

As an initial test, we screened KBM7 cells with a small library containing sgRNAs targeting the *BCR* and *ABL1* genes (table S3). The survival of KBM7 cells depends on the fusion protein produced by the BCR-ABL translocation (*26*). As expected, depletion was seen only for sgRNAs targeting the exons of *BCR* and *ABL1* that encode the fusion protein, but not for those targeting the other exons of *BCR* and *ABL1* (Fig. 3A).

We then infected Cas9-HL60, Cas9-KBM7, and WT KBM7 cells with the entire 73,000-member sgRNA library and used deep sequencing of the sgRNA barcodes to monitor the change in abundance of each sgRNA between the initial seeding and a final population obtained after 12 cell doublings (fig. S2, A and B).

We began by analyzing ribosomal protein genes, for which the library contained all possible sgRNAs. We observed strong Cas9-dependent depletion of sgRNAs targeting genes encoding ribosomal proteins, with good concordance between the sets of ribosomal protein genes essential for cell proliferation in the HL60 and KBM7 screens (the median sgRNA fold-change in abundance was used as a measure of gene essentiality) (Fig. 3, B and C). A few ribosomal protein genes were not found to be essential. These were two genes encoded on chromosome Y [*RPS4Y2*, which is testes-specific (*27*), and *RPS4YI,* which is expressed at low levels as compared with its homolog *RPS4X* on chromosome X (*28*)] and "ribosome-like" proteins, which may be required only in select tissues (*27*) and generally are lowly expressed in KBM7 cells (fig. S3A).

We then turned our attention to other genes within our data set, for which 10 sgRNAs were designed. As for the ribosomal genes, the essentiality scores of these genes were also strongly correlated between the two cells lines (fig. S3B and table S4). For the 20 highest scoring genes, we found independent evidence for essentiality, based primarily on data from large-scale functional studies in model organisms (table S5).

To evaluate the results at a global level, we tested 4722 gene sets to see whether they showed strong signatures of essentiality by using gene set enrichment analysis (*29*). Gene sets related to fundamental biological processes—including DNA replication, gene transcription, and protein degradation—showed strong depletion, which is consistent with their essentiality (Fig. 3D and table S6).

Last, we sought to understand the features underlying sgRNA efficacy. Although the vast majority of sgRNAs against ribosomal protein genes showed depletion, detailed comparison of sgRNAs targeting the same gene revealed substantial variation in the precise amounts of depletion. These differences are unlikely to be caused by local accessibility to the Cas9/sgRNA complex inasmuch as comparable variability was observed even among sgRNAs targeting neighboring target sites of a given gene (fig. S4A). Given that our library includes all possible sgRNAs against each of the 84 ribosomal genes, the data allowed us to search for factors that might explain the differential efficacy of sgRNAs. Because the majority of ribosomal protein genes are essential, we reasoned that the level of depletion of a given ribosomal protein-targeting sgRNA could serve as a proxy for its cleavage efficiency. Applying this approach, we found several trends related to sgRNA efficacy: (i) Single-guide sequences with very high or low GC content were less effective against their targets. (ii) sgRNAs targeting the last coding exon were less effective than those targeting earlier exons, which is consistent with the notion that disruption of the terminal exon would be expected to have less impact on gene function. (iii) sgRNAs targeting the transcribed strand were less effective than those targeting the nontranscribed strand (Fig. 3E). Although these trends were statistically significant, they explained only a small proportion of differences in sgRNA efficacy (table S7).

We hypothesized that differences in sgRNA efficacy might also result from sequence features governing interactions with Cas9. To test this, we developed a method to profile the sgRNAs directly bound to Cas9 in a highly parallel manner (supplementary materials, materials and methods).By comparing the abundance of sgRNAs bound to Cas9 relative to the abundance of their corresponding

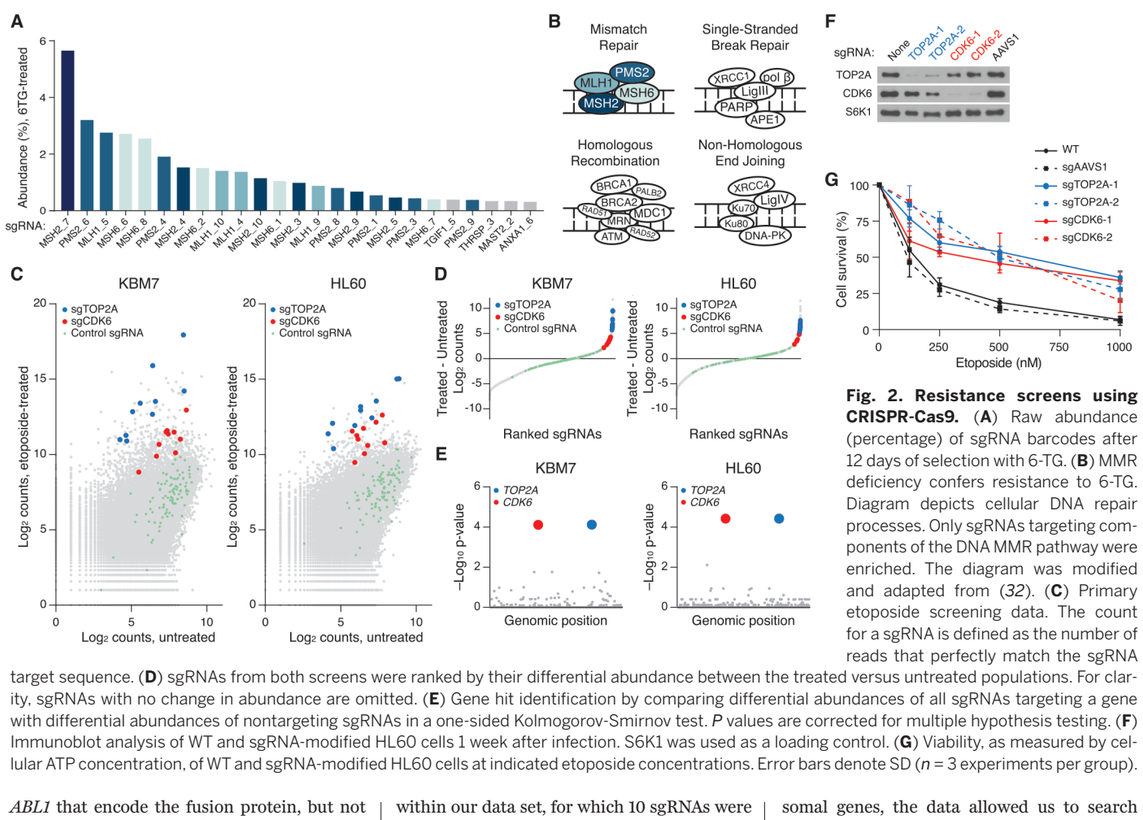


Fig. 2. Resistance screens using CRISPR-Cas9. (**A**) Raw abundance (percentage) of sgRNA barcodes after 12 days of selection with 6-TG. (**B**) MMR deficiency confers resistance to 6-TG. Diagram depicts cellular DNA repair processes. Only sgRNAs targeting components of the DNA MMR pathway were enriched. The diagram was modified and adapted from (*32*). (**C**) Primary etoposide screening data. The count for a sgRNA is defined as the number of reads that perfectly match the sgRNA target sequence. (**D**) sgRNAs from both screens were ranked by their differential abundance between the treated versus untreated populations. For clarity, sgRNAs with no change in abundance are omitted. (**E**) Gene hit identification by comparing differential abundances of all sgRNAs targeting a gene with differential abundances of nontargeting sgRNAs in a one-sided Kolmogorov-Smirnov test. *P* values are corrected for multiple hypothesis testing. (**F**) Immunoblot analysis of WT and sgRNA-modified HL60 cells 1 week after infection. S6K1 was used as a loading control. (**G**) Viability, as measured by cellular ATP concentration, of WT and sgRNA-modified HL60 cells at indicated etoposide concentrations. Error bars denote SD ($n = 3$ experiments per group).

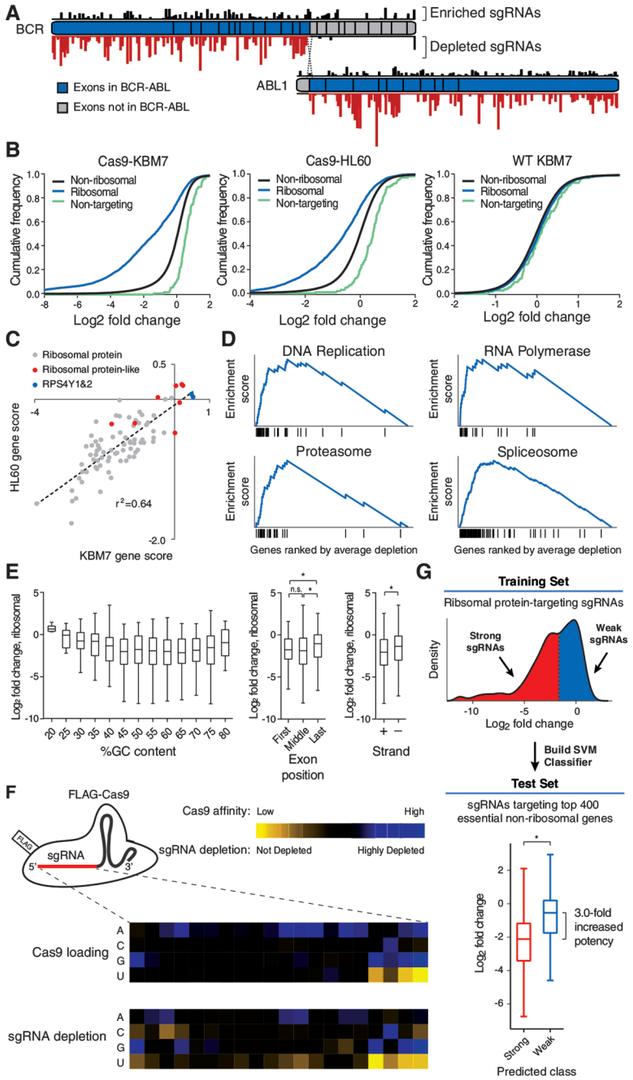


Fig. 3. Negative selection screens using CRISPR-Cas9 reveal rules governing sgRNA efficacy. (**A**) Selective depletion of sgRNAs targeting exons of *BCR* and *ABL1* present in the fusion protein. Individual sgRNAs are plotted according to their target sequence position along each gene, and the height of each bar indicates the level of depletion observed. Boxes indicate individual exons. (**B**) Cas9-dependent depletion of sgRNAs targeting ribosomal proteins. Cumulative distribution function plots of $\log_2$ fold changes in sgRNA abundance before and after 12 cell doublings in Cas9-KBM7, Cas9-HL60, and WT-KBM7 cells. (**C**) Requirement of similar sets of ribosomal protein genes for proliferation in the HL60 and KBM7 cells. Gene scores are defined as the median $\log_2$ fold change of all sgRNAs targeting a gene. (**D**) Depleted sgRNAs target genes involved in fundamental biological processes. Gene set enrichment analysis was performed on genes ranked by their combined depletion scores from screens in HL60 and KBM7 cells. Vertical lines underneath the *x* axis denote members of the gene set analyzed. (**E**) Features influencing sgRNA efficacy. Depletion ($\log_2$ fold change) of sgRNAs targeting ribosomal protein genes was used as an indicator of sgRNA efficacy. Correlation between $\log_2$ fold changes and spacer %GC content (left), exon position targeted (middle), and strand targeted (right) are depicted (*$P < 0.05$). (**F**) sgRNA target sequence preferences for Cas9 loading and cleavage efficiency. Position-specific nucleotide preferences for Cas9 loading are determined by counting sgRNAs bound to Cas9 normalized to the number of corresponding genomic integrations. Heatmaps depict sequence-dependent variation in Cas9 loading (top) and ribosomal protein gene-targeting sgRNA depletion (bottom). The color scale represents the median value (of Cas9 affinity or $\log_2$fold-change) for all sgRNAs with the specified nucleotide at the specified position. (**G**) sgRNA efficacy prediction. Ribosomal protein gene-targeting sgRNAs were designated as "weak" or "strong" on the basis of their $\log_2$ fold change and used to train a support-vector-machine (SVM) classifier. As an independent test, the SVM was used to predict the efficacy of sgRNAs targeting 400 essential nonribosomal genes (*$P < 0.05$).

genomic integrants, we found that the nucleotide composition near the 3′ end of the spacer sequence was the most important determinant of Cas9 loading (Fig. 3F). Specifically, Cas9 preferentially bound sgRNAs containing purines in the last four nucleotides of the spacer sequence, whereas pyrimidines were disfavored. A similar pattern emerged when we examined depletion of ribosomal protein-targeting sgRNAs [correlation coefficient (r) = 0.81], suggesting that, in significant part, the cleavage efficiency of a sgRNA was determined by its affinity for Cas9 (table S7).

We then sought to build an algorithm to discriminate between strong and weak sgRNAs (Fig. 3G). We trained a support-vector-machine classifier based on the target sequences and depletion scores of ribosomal protein-targeting sgRNAs. As an independent test, we used the classifier to predict the efficacy of sgRNAs targeting the 400 top scoring (essential) nonribosomal genes. The top two thirds of our predictions exhibited threefold higher efficacy than that of the remaining fraction, confirming the accuracy of the algorithm.

Using this algorithm, we designed a whole-genome sgRNA library consisting of sequences predicted to have higher efficacy (table S8). As with the sgRNA pool used in our screens, this new collection was also filtered for potential off-target matches. This reference set of sgRNAs may be useful both for targeting single genes as well as large-scale sgRNA screening.

Taken together, these results demonstrate the utility of CRISPR-Cas9 for conducting large-scale genetic screens in mammalian cells. On the basis of our initial experiments, this system appears to offer several powerful features that together provide substantial advantages over current functional screening methods.

First, CRISPR-Cas9 inactivates genes at the DNA level, making it possible to study phenotypes that require a complete loss of gene function to be elicited. In addition, the system should also enable functional interrogation of nontranscribed elements, which are inaccessible by means of RNAi.

Second, a large proportion of sgRNAs successfully generate mutations at their target sites. Although this parameter is difficult to directly assess in pooled screens, we can obtain an estimate by examining the "hit rate" at known genes. Applying a z score analysis of our positive selection screens, we found that over 75% (46 of 60) of sgRNAs score at a significance threshold that perfectly separates true and false positives on a gene level (fig. S5, A to D). Together, these results show that the effective coverage of our library is very high and that the rate of false negatives should be low, even in a large-scale screen.

Third, off-target effects do not appear to seriously hamper our screens, according to several lines of evidence. Direct sequencing of potential off-target loci detected minimal cleavage at secondary sites, which typically reside in noncoding regions and do not affect gene function. Moreover, in the 6-TG screens the 20 most abundant sgRNAs all targeted one of the four members of the MMR pathway. In total, they represented over 30% of the final pool, which is a fraction greater the next 400 sgRNAs combined. In the etoposide screen, the two top genes scored far above background levels (P values 100-fold smaller than that of the next best gene), enabling clear discrimination between true and false-positive hits. Last, new versions of the CRISPR-Cas9 system have recently been developed that substantially decrease off-target activity (*30, 31*).

Although we limited our investigation to proliferation-based phenotypes, our approach can be applied to a much wider range of biological phenomena. With appropriate sgRNA libraries, the method should enable genetic analyses of mammalian cells to be conducted with a degree of rigor and completeness currently possible only in the study of microorganisms.

REFERENCES AND NOTES

1. G. Giaever *et al.*, *Nature* **418**, 387–391 (2002).
2. M. Costanzo *et al.*, *Science* **327,** 425–431 (2010).
3. J. E. Carette *et al.*, *Science* **326**, 1231–1235 (2009).
4. G. Guo, W. Wang, A. Bradley, *Nature* **429**, 891–895 (2004).
5. A. Fire *et al.*, *Nature* **391**, 806–811 *et al.*, (1998).
6. T. R. Brummelkamp, R. Bernards, R. Agami, *Science* **296**, 550–553 (2002).
7. J. Moffat *et al.*, *Cell* **124**, 1283–1298 (2006).
8. V. N. Ngo *et al.*, *Nature* **441**, 106–110 (2006).
9. H. W. Cheung *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 12372–12377 (2011).
10. C. J. Echeverri *et al.*, *Nat. Methods* **3**, 777–779 (2006).
11. M. Booker *et al.*, *BMC Genomics* **12**, 50 (2011).
12. W. G. Kaelin Jr., *Science* **337**, 421–422 (2012).
13. R. Barrangou *et al.*, *Science* **315**, 1709–1712 (2007).
14. L. Cong *et al.*, *Science* **339**, 819–823 (2013).
15. P. Mali *et al.*, *Science* **339**, 823–826 (2013).
16. W. Y. Hwang *et al.*, *Nat. Biotechnol.* **31**, 227–229 (2013).
17. M. Jinek *et al.*, *Science* **337**, 816–821 (2012).
18. T. Horii, D. Tamura, S. Morita, M. Kimura, I. Hatada, *Int. J. Mol. Sci.* **14**, 19774–19781 (2013).
19. H. Wang *et al.*, *Cell* **153**, 910–918 (2013).
20. P. D. Hsu *et al.*, *Nat. Biotechnol.* **31**, 827–832 (2013).
21. Y. Fu *et al.*, *Nat. Biotechnol.* **31**, 822–826 (2013).
22. V. Pattanayak *et al.*, *Nat. Biotechnol.* **31**, 839–843 (2013).
23. T. Yan, S. E. Berry, A. B. Desai, T. J. Kinsella, *Clin. Cancer Res.* **9**, 2327–2334 (2003).
24. R. D. Kolodner, G. T. Marsischky, *Curr. Opin. Genet. Dev.* **9**, 89–96 (1999).
25. D. J. Burgess *et al.*, *Proc. Natl. Acad.* Sci. *U.S.A.* **105**, 9053–9058 (2008).
26. B. Scappini *et al.*, *Cancer* **100**, 1459–1471 (2004).
27. S. Xue, M. Barna, *Nat. Rev. Mol. Cell Biol.* **13**, 355–369 (2012).
28. C. M. Johnston *et al.*, *PLOS Genet.* **4**, e9 (2008).
29. A. Subramanian *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
30. F. A. Ran *et al.*, *Cell* **154**, 1380–1389 (2013).
31. P. Mali *et al.*, *Nat. Biotechnol.* **31**, 833–838 (2013).
32. K. Yoshimoto *et al.*, *Front. Oncol.* **2**, 186 (2012).

ACKNOWLEDGMENTS

We thank all members of the Sabatini and Lander labs, especially J. Engreitz, S. Schwartz, A. Shishkin, and Z. Tsun for protocols, reagents, and advice; T. Mikkelsen for assistance with oligonucleotide synthesis; and L. Gaffney for assistance with figures. This work was supported by the U.S. National Institutes of Health (CA103866) (D.M.S.), National Human Genome Research Institute (2U54HG003067-10) (E.S.L.), the Broad Institute of MIT and Harvard (E.S.L.), and an award from the U.S. National Science Foundation (T.W.). The composition of the sgRNA pools and screening data can be found in the supplementary materials. A patent application has been filed by the Broad Institute relating to aspects of the work described in this manuscript. Inducible Cas9 and sgRNA backbone lentiviral vectors and the genome-scale sgRNA plasmid pool are deposited in Addgene.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/343/6166/80/suppl/DC1
Materials and Methods Supplementary Text
Figs. S1 to S5
Tables S1 to S8
References (*33–43*)

8 October 2013; accepted 2 December 2013
Published online 12 December 2013;
10.1126/science.1246981

Genome-scale CRISPR-Cas9 knockout screening in human cells

Ophir Shalem,[1,2*] Neville E. Sanjana,[1,2*] Ella Hartenian,[1] Xi Shi,[1,3] David A. Scott,[1,2] Tarjei S. Mikkelsen,[1] Dirk Heckl,[4] Benjamin L. Ebert,[4] David E. Root,[1] John G. Doench,[1] Feng Zhang[1,2†]

The simplicity of programming the CRISPR (clustered regularly interspaced short palindromic repeats)–associated nuclease Cas9 to modify specific genomic loci suggests a new way to interrogate gene function on a genome-wide scale. We show that lentiviral delivery of a genome-scale CRISPR-Cas9 knockout (GeCKO) library targeting 18,080 genes with 64,751 unique guide sequences enables both negative and positive selection screening in human cells. First, we used the GeCKO library to identify genes essential for cell viability in cancer and pluripotent stem cells. Next, in a melanoma model, we screened for genes whose loss is involved in resistance to vemurafenib, a therapeutic RAF inhibitor. Our highest-ranking candidates include previously validated genes *NF1* and *MED12*, as well as novel hits *NF2*, *CUL3*, *TADA2*B, and *TADA1*. We observe a high level of consistency between independent guide RNAs targeting the same gene and a high rate of hit confirmation, demonstrating the promise of genome-scale screening with Cas9.

A major goal since the completion of the Human Genome Project is the functional characterization of all annotated genetic elements in normal biological processes and disease (*1*). Genome-scale loss-of-function screens have provided a wealth of information in diverse model systems (*2–5*). In mammalian cells, RNA interference (RNAi) is the predominant method for genome-wide loss-of-function screening (*2, 3*), but its utility is limited by the inherent incompleteness of protein depletion by RNAi and confounding off-target effects (*6, 7*).

The RNA-guided CRISPR (clustered regularly interspaced short palindrome repeats)–associated nuclease Cas9 provides an effective means of introducing targeted loss-of-function mutations at specific sites in the genome (*8, 9*). Cas9 can be programmed to induce DNA double-strand breaks (DSBs) at specific genomic loci (*8, 9*) through a synthetic single-guide RNA (sgRNA) (*10*), which when targeted to coding regions of genes can create frame shift insertion/deletion (indel) mutations that result in a loss-of-function allele. Because the targeting specificity of Cas9 is conferred by short guide sequences, which can be easily generated at large scale by array-based oligonucleotide library synthesis (*11*), we sought to explore the potential of Cas9 for pooled genome-scale functional screening.

Lentiviral vectors are commonly used for delivery of pooled short-hairpin RNAs (shRNAs) in RNAi because they can be easily titrated to control transgene copy number and are stably maintained as genomic integrants during subsequent cell replication

[1]Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA.
[2]McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
[3]Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, 7 Cambridge Center, Cambridge, MA 02142, USA.
[4]Division of Hematology, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA.
*These authors contributed equally to this work.
†Corresponding author. E-mail: zhang@broadinstitute.org

(*2*, *12*, *13*). Therefore, we designed a single lentiviral vector to deliver Cas9, a sgRNA, and a puromycin selection marker into target cells (lentiCRISPR) (Fig. 1A). The ability to simultaneously deliver Cas9 and sgRNA through a single vector enables application to any cell type of interest, without the need to first generate cell lines that express Cas9.

To determine the efficacy of gene knockout by lentiCRISPR transduction, we tested six sgRNAs targeting enhanced green fluorescent protein (EGFP) in a human embryonic kidney (HEK) 293T cell line containing a single copy of EGFP (fig. S1). After transduction at a low multiplicity of infection (MOI = 0.3) followed by selection with puromycin, lentiCRISPRs abolished EGFP fluorescence in 93 ± 8% (mean ± SD) of cells after 11 days (Fig. 1B). Deep sequencing of the EGFP locus revealed a 92 ± 9% indel frequency ($n \geq 10^4$ sequencing reads per condition) (fig. S2). In contrast, transduction of cells with lentiviral vectors expressing EGFP-targeting shRNA led to incomplete knockdown of EGFP fluorescence (Fig. 1C).

Given the high efficacy of gene knockout by lentiCRISPR, we tested the feasibility of conducting genome-scale CRISPR-Cas9 knockout (GeCKO) screening with a pooled lentiCRISPR library. We designed a library of sgRNAs targeting 5′ constitutive exons (Fig. 2A) of 18,080 genes in the human genome with an average coverage of 3 to 4 sgRNAs per gene (table S1), and each target site was selected to minimize off-target modification (*14*) (see supplementary text).

To test the efficacy of the full GeCKO library at achieving knockout of endogenous gene targets, we conducted a negative selection screen by profiling the depletion of sgRNAs targeting essential survival genes (Fig. 2A). We transduced the human melanoma cell line A375 and the human stem cell line HUES62 with the GeCKO library at a MOI of 0.3. As expected, deep sequencing (figs. S3 and S4) 14 days after transduction revealed a significant reduction in the diversity of sgRNAs in the surviving A375 and HUES62 cells (Fig. 2, B and C) (Wilcoxon rank sum test, $P < 10^{-10}$ for both cell types). Gene set enrichment analysis (GSEA) (*15*) indicated that most of the depleted sgRNAs targeted essential genes such as ribosomal structural constituents (Fig. 2, D and E, and tables S2 and S3). The overlap in highly depleted genes and functional gene categories between the two cell lines (fig. S5) indicates that GeCKO can identify essential genes and that enrichment analysis of depleted sgRNAs can pinpoint key functional gene categories in negative selection screens.

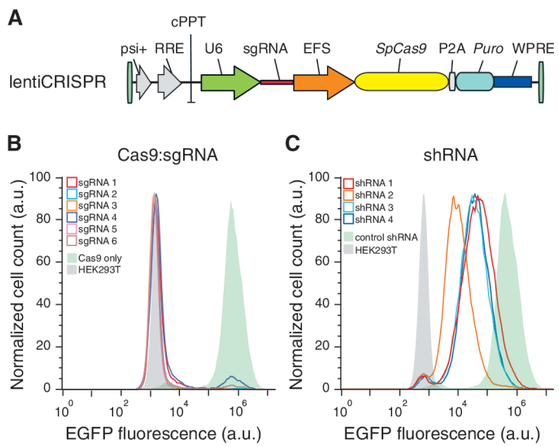


Fig. 1. Lentiviral delivery of Cas9 and sgRNA provides efficient depletion of target genes. (**A**) Lentiviral expression vector for Cas9 and sgRNA (lentiCRISPR). puro, puromycin selection marker; psi+, psi packaging signal; RRE, *rev* response element; cPPT, central polypurine tract; EFS, elongation factor-1α short promoter; P2A, 2A self-cleaving peptide; WPRE, posttranscriptional regulatory element. (**B**) Distribution of fluorescence from 293T-EGFP cells transduced by EGFP-targeting lentiCRISPR (sgRNAs 1 to 6, outlined peaks) and Cas9-only (green-shaded peak) vectors, and nonfluorescent 293T cells (gray shaded peak). (**C**) Distribution of fluorescence from 293T-EGFP cells transduced by EGFP-targeting shRNA (shRNAs 1 to 4, outlined peaks) and control shRNA (green-shaded peak) vectors, and nonfluorescent 293T cells (gray shaded peak).

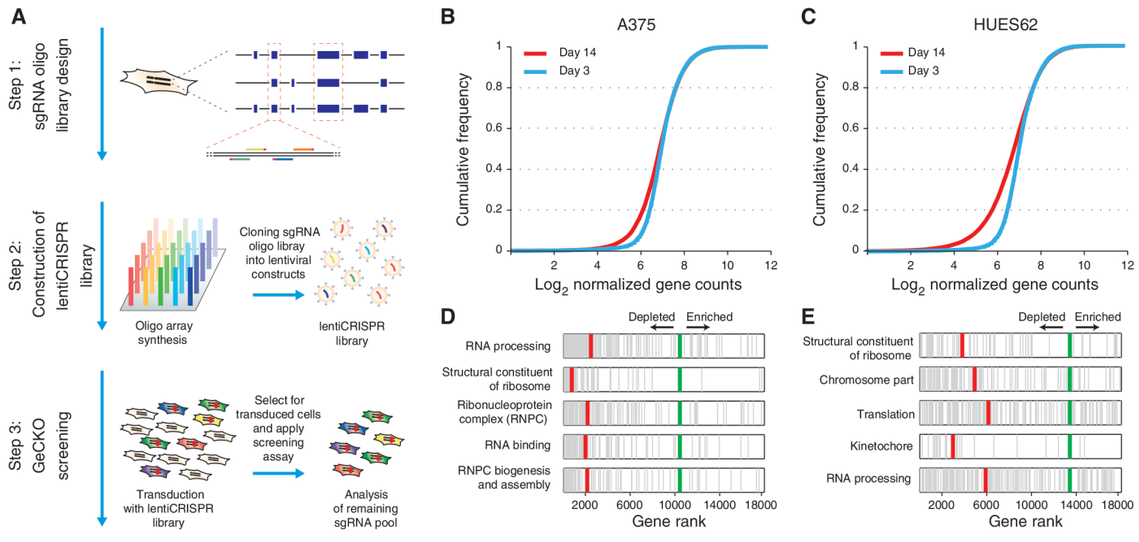


Fig. 2. GeCKO library design and application for genome-scale negative selection screening. (**A**) Design of sgRNA library for genome-scale knockout of coding sequences in human cells (see supplementary text). (**B** and **C**) Cumulative frequency of sgRNAs 3 and 14 days after transduction in A375 and human embryonic stem cells, respectively. Shift in the 14-day curve represents the depletion in a subset of sgRNAs. (**D** and **E**) The five most significantly depleted gene sets in A375 cells [nominal $P < 10^{-5}$, false discovery rate (FDR)–corrected $q < 10^{-5}$] and HUES62 cells (nominal $P < 10^{-5}$, FDR-corrected $q < 10^{-3}$) identified by GSEA (*15*).

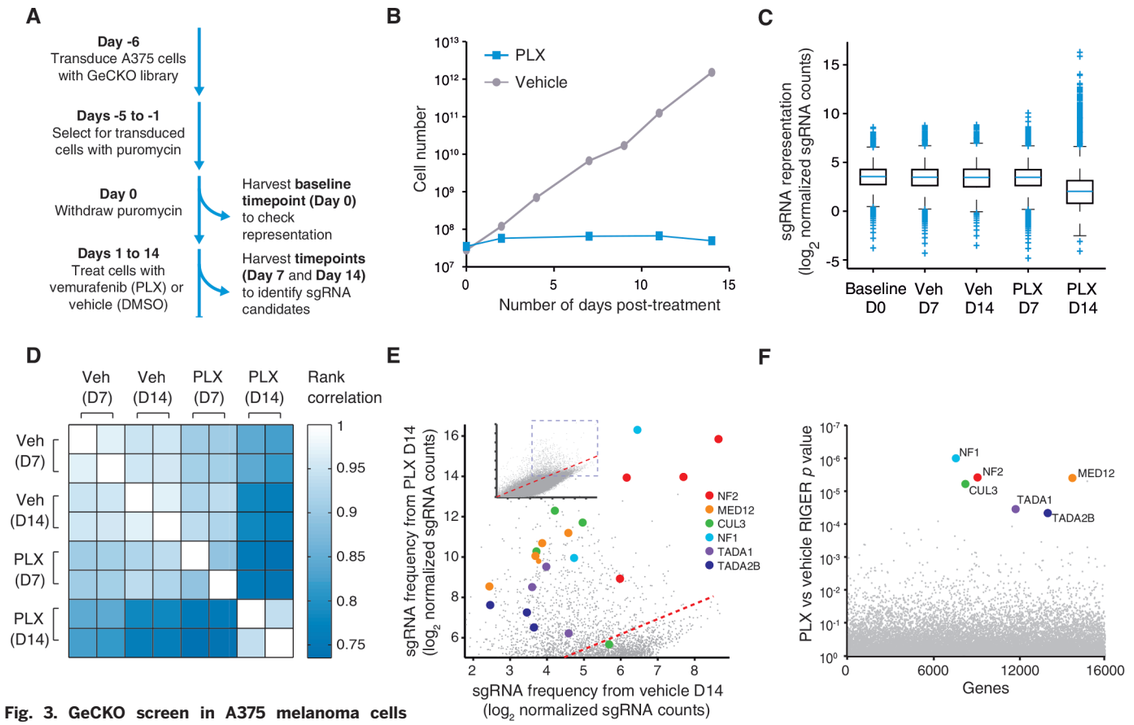


Fig. 3. GeCKO screen in A375 melanoma cells reveals genes whose loss confers PLX resistance. (**A**) Timeline of PLX resistance screen in A375 melanoma cells. (**B**) Growth of A375 cells when treated with dimethyl sulfoxide (DMSO) or PLX over 14 days. (**C**) Box plot showing the distribution of sgRNA frequencies at different time points, with and without PLX treatment (vehicle is DMSO). The box extends from the first to the third quartile with the whiskers denoting 1.5 times the interquartile range. Enrichment of specific sgRNAs: 7 days of PLX treatment, 1 sgRNA greater than 10-fold enrichment; 14 days of PLX treatment, 379 and 49 sgRNAs greater than 10-fold and 100-fold enrichment, respectively. (**D**) Rank correlation of normalized sgRNA read count between biological replicates and treatment conditions. (**E**) Scatterplot showing enrichment of specific sgRNAs after PLX treatment. (**F**) Identification of top candidate genes using the RIGER *P* value analysis.

For a subset of genes, we found enrichment of multiple sgRNAs that target each gene after 14 days of PLX treatment (Fig. 3E), suggesting that loss of these particular genes contributes to PLX resistance. We used the RNAi Gene Enrichment Ranking (RIGER) algorithm to rank screening hits by the consistent enrichment among multiple sgRNAs targeting the same gene (Fig. 3F and table S4) (*12*). Our highest-ranking genes included previously reported candidates *NF1* and *MED12* (*18*, *19*) and also several genes not previously implicated in PLX resistance, including neurofibromin 2 (*NF2*), Cullin 3 E3 ligase (*CUL3*), and members of the STAGA histone acetyltransferase complex (*TADA1* and *TADA2B*). These candidates yield new testable hypotheses regarding PLX resistance mechanisms (see supplementary text). For example, *NF1* and *NF2*, although unrelated in sequence, are each mutated in similar but distinct forms of neurofibromatosis (*20*). In addition, epigenetic dysregulation resulting from mutations in the mechanistically related STAGA and Mediator complexes (*21*) may have a role in acquired drug resistance. All of these hits were also identified through a second independent library transduction (figs. S7 and S8 and tables S5 and S6).

A similar screen to identify PLX drug resistance in A375 cells was previously conducted using a pooled library of 90,000 shRNAs (*19*). To compare the efficacy and reliability of genome-scale shRNA screening with GeCKO, we used several methods to evaluate the degree of consistency among the sgRNAs or shRNAs targeting the top candidate genes. First, we plotted the *P* values for the top 100 hits using either RIGER (Fig. 4A) or redundant siRNA activity (RSA) (fig. S9) scoring. Lower *P* values for the GeCKO versus shRNA screen indicate better scoring consistency among sgRNAs. Second, for the top 10 RIGER hit genes, 78 ± 27% of sgRNAs targeting each gene ranked among the top 5% of enriched sgRNAs, whereas 20 ± 12% of shRNAs targeting each gene ranked among the top 5% of enriched shRNAs (Fig. 4B).

We validated top-ranking genes from the GeCKO screen individually using 3 to 5 sgRNAs (Fig. 4, C to E, and figs. S10 and S11).

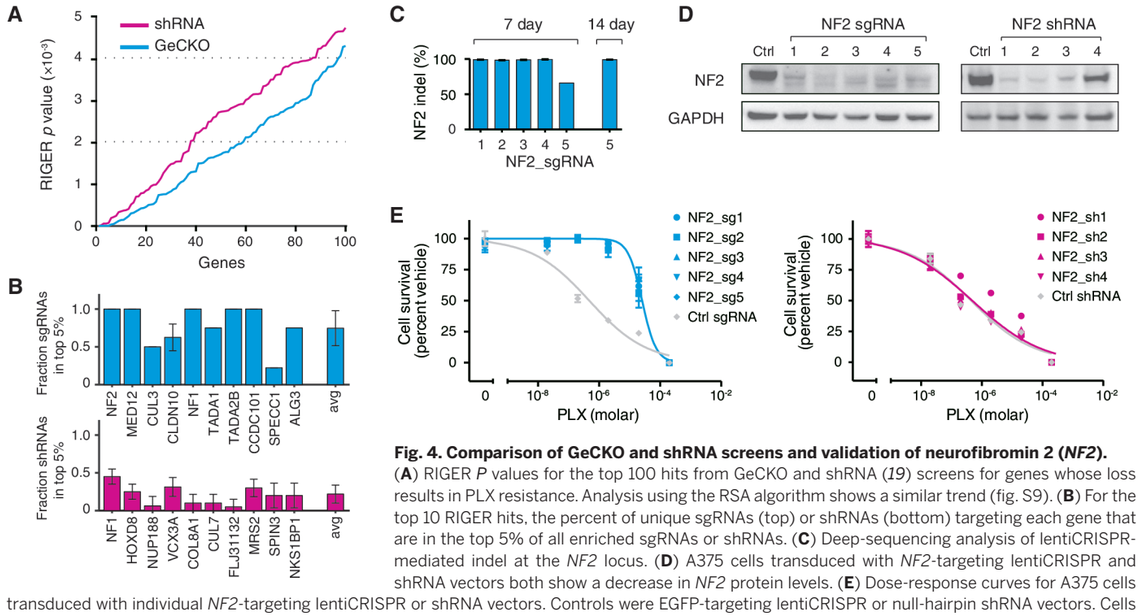


Fig. 4. Comparison of GeCKO and shRNA screens and validation of neurofibromin 2 (*NF2*). (**A**) RIGER *P* values for the top 100 hits from GeCKO and shRNA (*19*) screens for genes whose loss results in PLX resistance. Analysis using the RSA algorithm shows a similar trend (fig. S9). (**B**) For the top 10 RIGER hits, the percent of unique sgRNAs (top) or shRNAs (bottom) targeting each gene that are in the top 5% of all enriched sgRNAs or shRNAs. (**C**) Deep-sequencing analysis of lentiCRISPR-mediated indel at the *NF2* locus. (**D**) A375 cells transduced with *NF2*-targeting lentiCRISPR and shRNA vectors both show a decrease in *NF2* protein levels. (**E**) Dose-response curves for A375 cells transduced with individual *NF2*-targeting lentiCRISPR or shRNA vectors. Controls were EGFP-targeting lentiCRISPR or null-hairpin shRNA vectors. Cells transduced with *NF2*-targeting lentiCRISPRs show a significant increase ($F_{1,8}$ = 30.3, P < 0.001, n = 4 replicates) in the half-maximal effective concentration (EC_{50}), whereas cells transduced with *NF2*-targeting shRNA vectors do not ($F_{1,8}$ = 0.47, P = 0.51, n = 4 replicates).

For *NF2*, we found that 4 out of 5 sgRNAs resulted in >98% allele modification 7 days after transduction, and all 5 sgRNAs showed >99% allele modification 14 days after transduction (Fig. 4C). We compared sgRNA and shRNA-mediated protein depletion and PLX resistance using Western blot (Fig. 4D) and cell growth assays (Fig. 4E). Interestingly, although all five sgRNAs conferred resistance to PLX, only the best shRNA achieved sufficient knockdown to increase PLX resistance (Fig. 4E), suggesting that even low levels of *NF2* are sufficient to retain sensitivity to PLX. Additionally, sgRNAs targeting *NF1*, *MED12*, *CUL3*, *TADA1*, and *TADA2B* led to a decrease in protein expression and increased resistance to PLX (figs. S10 and S11). Deep sequencing confirmed a high rate of mutagenesis at targeted loci (figs. S12 and S13), with a small subset of off-target sites exhibiting indels (figs. S14 to S16), which may be alleviated using an offset nicking approach (*22, 23*) that was recently shown to reduce off-target modifications (*22*).

GeCKO screening provides a mechanistically distinct method from RNAi for systematic perturbation of gene function. Whereas RNAi reduces protein expression by targeting RNA, GeCKO introduces loss-of-function mutations into genomic DNA. Although some indel mutations are expected to maintain the reading frame, homozygous knockout yields high screening sensitivity, which is especially important in cases where incomplete knockdown retains gene function. In addition, RNAi is limited to transcripts, whereas Cas9:sgRNAs can target elements across the entire genome, including promoters, enhancers, introns, and intergenic regions. Furthermore, catalytically inactive mutants of Cas9 can be tethered to different functional domains (*23–27*) to broaden the repertoire of perturbation modalities, including genome-scale gain-of-function screening using Cas9 activators and epigenetic modifiers. In the GeCKO screens presented here, the efficiency of complete knockout, the consistency of distinct sgRNAs, and the high validation rate for top screen hits demonstrate the potential of Cas9:sgRNA-based technology to transform functional genomics.

REFERENCES AND NOTES

1. E. S. Lander, *Nature* **470**, 187–197 (2011).
2. K. Berns *et al.*, *Nature* **428**, 431–437 (2004).
3. M. Boutros *et al.*, Heidelberg Fly Array Consortium, *Science* **303**, 832–835 (2004).
4. R. Rad *et al.*, *Science* **330**, 1104–1107 (2010).
5. J. E. Carette *et al.*, *Science* **326**, 1231–1235 (2009).
6. A. L. Jackson *et al.*, *RNA* **12**, 1179–1187 (2006).
7. C. J. Echeverri *et al.*, *Nat. Methods* **3**, 777–779 (2006).
8. L. Cong *et al.*, *Science* **339**, 819–823 (2013).
9. P. Mali *et al.*, *Science* **339**, 823–826 (2013).
10. M. Jinek *et al.*, *Science* **337**, 816–821 (2012).
11. A. P. Blanchard, L. Hood, *Nat. Biotechnol.* **14**, 1649 (1996).
12. B. Luo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 20380–20385 (2008).
13. P. J. Paddison *et al.*, *Nature* **428**, 427–431 (2004).
14. P. D. Hsu *et al.*, *Nat. Biotechnol.* **31**, 827–832 (2013).
15. A. Subramanian *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
16. K. T. Flaherty *et al.*, *N. Engl. J. Med.* **363**, 809–819 (2010).
17. H. Davies *et al.*, *Nature* **417**, 949–954 (2002).
18. S. Huang *et al.*, *Cell* **151**, 937–950 (2012).
19. S. R. Whittaker *et al.*, *Cancer Discovery* **3**, 350–362 (2013).
20. A. L. Lin, D. H. Gutmann, *Nat. Rev. Clin. Oncol.* **10**, 616–624 (2013).
21. X. Liu, M. Vorontchikhina, Y. L. Wang, F. Faiola E. Martinez, *Mol. Cell. Biol.* **28**, 108–121 (2008).
22. F. A. Ran *et al.*, *Cell* **154**, 1380–1389 (2013).
23. P. Mali *et al.*, Nat. *Biotechnol.* **31**, 833–838 (2013).
24. L. A. Gilbert *et al.*, *Cell* **154**, 442–451 (2013).
25. S. Konermann *et al.*, *Nature* **500**, 472–476 (2013).
26. P. Perez-Pinera *et al.*, *Nat. Methods* **10**, 973–976 (2013).
27. M. L. Maeder *et al.*, *Nat. Methods* **10**, 977–979 (2013).

ACKNOWLEDGMENTS

We thank G. Cowley, W. Harrington, J. Wright, E. Hodis, S. Whittaker, J. Merkin, C. Burge, D. Peters, C. Cowan, L. P. Club, and the entire Zhang laboratory for technical support and critical discussions. O.S. is a Klarman Family Foundation Fellow, N.S. is a Simons Center for the Social Brain Postdoctoral Fellow, D.A.S. is an NSF Fellow, and J.D. is a Merkin Institute Fellow. D. H. was funded by the German Cancer Foundation. F.Z. is supported by an NIH Director's Pioneer Award (1DP1-MH100706); a NIH Transformative R01 grant (1R01-DK097768); the Keck, McKnight, Merkin, Vallee, Damon Runyon, Searle Scholars, Klingenstein, and Simons Foundations; and Bob Metcalfe and Jane Pauley. The authors have no conflicting financial interests. A patent application has been filed relating to this work, and the authors plan to make the reagents widely available to the academic community through Addgene and to provide software tools at the Zhang laboratory Web site (www.genome-engineering.org).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/343/6166/84/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S16
Tables S1 to S10
References

9 October 2013; accepted 2 December 2013
Published online 12 December 2013;
10.1126/science.1247005

The CRISPR craze

Elizabeth Pennisi

A bacterial immune system yields a potentially revolutionary genome-editing technique.

Bacteria may not elicit much sympathy from us eukaryotes, but they, too, can get sick. That's potentially a big problem for the dairy industry, which often depends on bacteria such as *Streptococcus thermophilus* to make yogurts and cheeses. *S. thermophilus* breaks down the milk sugar lactose into tangy lactic acid. But certain viruses—bacteriophages, or simply phages—can debilitate the bacterium, wreaking havoc on the quality or quantity of the food it helps produce.

In 2007, scientists from Danisco, a Copenhagen-based food ingredient company now owned by DuPont, found a way to boost the phage defenses of this workhouse microbe. They exposed the bacterium to a phage and showed that this essentially vaccinated it against that virus (*Science*, 23 March 2007, p. 1650). The trick has enabled DuPont to create heartier bacterial strains for food production. It also revealed something fundamental: Bacteria have a kind of adaptive immune system, which enables them to fight off repeated attacks by specific phages.

That immune system has suddenly become important for more than food scientists and microbiologists, because of a valuable feature: It takes aim at specific DNA sequences. In January, four research teams reported harnessing the system, called CRISPR for peculiar features in the DNA of bacteria that deploy it, to target the destruction of specific genes in human cells. And in the following 8 months, various groups have used it to delete, add, activate, or suppress targeted genes in human cells, mice, rats, zebrafish, bacteria, fruit flies, yeast, nematodes, and crops, demonstrating broad utility for the technique. Biologists had recently developed several new ways to precisely manipulate genes, but CRISPR's "efficiency and ease of use trumps just about anything," says George Church of Harvard University, whose lab was among the first to show that the technique worked in human cells.

With CRISPR, scientists can create mouse models of human diseases much more quickly than before, study individual genes much faster, and easily change multiple genes in cells at once to study their interactions. This year's CRISPR craze may yet slow down as limitations of the method emerge, but Church and other CRISPR pioneers are already forming companies to harness the technology for treating genetic diseases. "I don't think there's any example of any field moving this fast," says Blake Wiedenheft, a biochemist at Montana State University in Bozeman.

PHOTOS: (TOP TO BOTTOM) REPRINTED BY PERMISSION FROM MACMILLAN PUBLISHERS LTD: *NATURE METHODS* (AE FRIEDLAND *ET AL.*, **10**, 741–743 (2013) DOI:10.1038/NMETH.2532); GONZALES, A.P.W., YEH, J.-R. & PETERSON, R.T.; SCOTT GRATZ/UNIVERSITY OF WISCONSIN, MADISON.

Precise cuts. In just 8 months, CRISPR modifications of DNA resulted in dumpier nematodes (*top*, bottom), zebrafish embryos with an excess of ventral tissue (*middle*, bottom), and fruit flies with dark eyes (*bottom*, right), demonstrating its broad utility for editing genes in animals.

Humble beginnings

The first inkling of this hot new genetic engineering tool came in 1987, when a research team observed an oddly repetitive sequence at one end of a bacterial gene. Few others took much notice. A decade later, though, biologists deciphering microbial genomes often found similar puzzling patterns, in which a sequence of DNA would be followed by nearly the same sequence in reverse, then 30 or so seemingly random bases of "spacer DNA," and then a repeat of the same palindromic sequence, followed by a different spacer DNA. A single microbe could have several such stretches, each with different repeat and intervening sequences. This pattern appears in more than 40% of bacteria and fully 90% of microbes in a different domain, the archaea, and gives CRISPR its name. (It stands for clustered regularly interspaced short palindromic repeats.)

Many researchers assumed that these odd sequences were junk, but in 2005, three bioinformatics groups reported that spacer DNA often matched the sequences of phages, indicating a possible role for CRISPR in microbial immunity. "That was a very key clue," says biochemist Jennifer Doudna of the University of California (UC), Berkeley. It led Eugene Koonin from the National Center for Biotechnology Information in Bethesda, Maryland, and his colleagues to propose that bacteria and archaea take up phage DNA, then preserve it as a template for molecules of RNA that can stop matching foreign DNA in its tracks, much the way eukaryotic cells use a system called RNA interference (RNAi) to destroy RNA.

Enter the Danisco team. In 2007, Rodolphe Barrangou, Philippe Horvath, and others with the company showed that they could alter the resistance of *S. thermophilus* to phage attack by adding or deleting spacer DNA that matched the phage's. At the time, Barrangou, who is now at North Carolina State University in Raleigh, didn't see CRISPR's full potential. "We had no idea that those elements could be readily exploitable for something as attractive as genome editing," he says.

Doudna and Emmanuelle Charpentier, currently of the Helmholtz Centre for Infection Research and Hannover Medical School in Germany, took the next step. They had independently been teasing out the roles of various CRISPR-associated proteins to learn how bacteria deploy the DNA spacers in their immune defenses. But the duo soon joined forces to focus on a CRISPR system that relies on a protein called Cas9, as it was simpler than other CRISPR systems.

When CRISPR goes into action in response to an invading phage, bacteria transcribe the spacers and the palindromic DNA into a long RNA molecule that the cell then cuts into short spacer-derived RNAs called crRNAs. An additional stretch of RNA, called tracrRNA, works with Cas9 to produce the crRNA, Charpentier's group reported in *Nature* in 2011. The group proposed that together, Cas9, tracrRNA, and crRNA somehow attack foreign DNA that matches the crRNA.

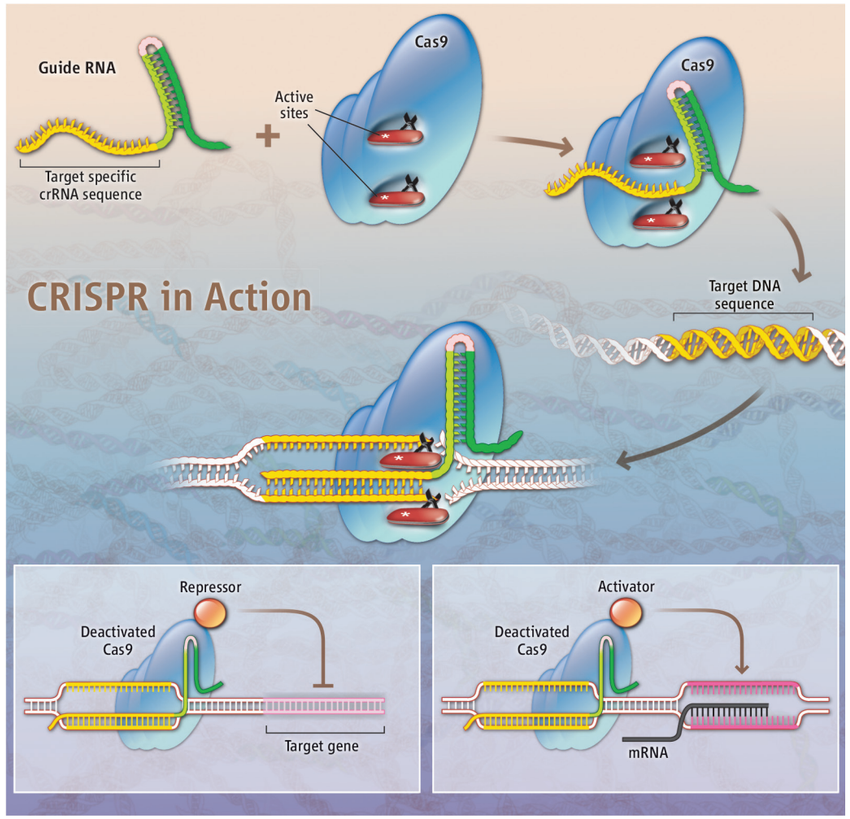
The two teams found that the Cas9 protein is a nuclease, an enzyme specialized for cutting DNA, with two active cutting sites,

one site for each strand of the DNA's double helix. And in a discovery that foreshadowed CRISPR's broad potential for genome engineering, the team demonstrated that they could disable one or both cutting sites without interfering with the ability of the complex to home in on its target DNA. "The possibility of using a single enzyme by just changing the RNA seemed very simple," Doudna recalls.

Before CRISPR could be put to use, however, Doudna's and Charpentier's teams had to show that they could control where Cas9 went to do its cutting. First, Doudna's postdoc, Martin Jinek, figured out how to combine tracrRNA and spacer RNA into a "single-guide RNA" molecule; then, as a proof of principle, the team last year made several guide RNAs, mixed them with Cas9, and showed in a test tube that the synthetic complexes could find and cut their DNA targets (*Science*, 17 August 2012, p. 816). "That was a milestone paper," Barrangou says.

This precision targeting drives the growing interest in CRISPR. Genetic engineers have long been able to add and delete genes in a number of organisms. But they couldn't dictate where those genes would insert into the genome or control where gene deletions occurred. Then, a decade ago, researchers developed zinc finger nucleases, synthetic proteins that have DNA-binding domains that enable them to home in and break DNA at specific spots. A welcome addition to the genetic engineering toolbox, zinc fingers even spawned a company that is testing a zinc finger to treat people infected with HIV (*Science*, 23 December 2005, p. 1894). More recently, synthetic nucleases called TALENs have proved an easier way to target specific DNA and were predicted to surpass zinc fingers (*Science*, 14 December 2012, p. 1408).

Now, CRISPR systems have stormed onto the scene, promising to even outcompete TALENs. Unlike the CRISPR system, which uses RNA as its DNA-homing mechanism, zinc finger and TALEN technologies both depend on custom-making new proteins for each DNA target. The CRISPR system's "guide RNAs" are much easier to make than proteins, Barrangou says. "Within a couple weeks you can generate very tangible results that using alternative methods would take months."



DNA surgeon. With just a guide RNA and a protein called Cas9, researchers first showed that the CRISPR system can home in on and cut specific DNA, knocking out a gene or enabling part of it to be replaced by substitute DNA. More recently, Cas9 modifications have made possible the repression (*lower left*) or activation (*lower right*) of specific genes.

ILLUSTRATION: K. SUTLIFF/*SCIENCE*

Harnessing CRISPR

Speed is not its only advantage. Church's group had been pushing the use of TALENs in human cells, but when he learned of Doudna and Charpentier's results, he and his colleagues made guide RNA against genes they had already targeted with TALENs. In three human cell types, the CRISPR system was more efficient than TALENs at cutting the DNA target, and it worked on more genes than TALENs did (*Science*, 15 February 2013, p. 823). To demonstrate the ease of the CRISPR system, Church's team synthesized a library of tens of thousands of guide RNA sequences, capable of targeting 90% of human genes. "You can pepper the genome with every imaginable CRISPR," he says.

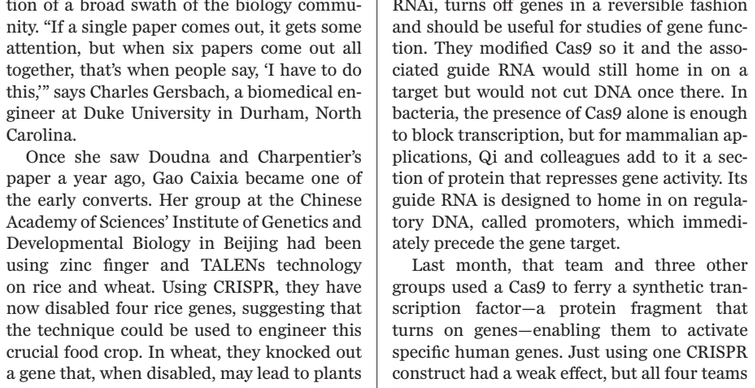
That makes it possible to alter virtually any gene with Cas9, exploiting its DNA-cutting ability to either disable the gene or cut it apart, allowing substitute DNA to be inserted. In an independent paper that appeared at the same time as Church's, Feng Zhang, a synthetic biologist at the Broad Institute in Cambridge, Massachusetts, and his colleagues showed that CRISPR can target and cut two genes at once in human cells (*Science*, 15 February 2013, p. 819). And working with developmental biologist Rudolf Jaenisch at the Whitehead Institute for Biomedical Research in Cambridge, Zhang has since disrupted five genes at once in mouse embryonic stem (ES) cells.

Such work lays the foundation for generating mutant mice, a key tool for biomedical research. One approach would be to add the altered mouse ES cells to a developing embryo and breed the resulting animals. But Zhang has demonstrated a faster option. His team found it could simply inject fertilized mouse eggs, or zygotes, with Cas9 messenger RNA and two guide RNAs and, with 80% efficiency, knock out two genes. They could also perform more delicate genomic surgery on the embryos by shackling Cas9, so that it nicks target DNA instead of cutting it. In this way, they could introduce a new part of a gene through a process called homology-directed repair, they reported in the 2 May 2013 issue of *Cell*.

Developing a new mouse model for a disease now entails careful breeding of multiple generations and can take a year; with Zhang's CRISPR technique, a new mouse model could be ready for testing in a matter of weeks. And Zhang thinks the approach is not limited to mice. "As long as you can manipulate the embryo and then reimplant it, then you will be able to do it" in larger animals, perhaps even primates.

Doudna's group and a Korean team reported using CRISPR to cut DNA in human cells 3 weeks after Zhang's and Church's papers went online, and, at the same time, another group revealed they had used CRISPR to make mutant zebrafish. This cascade of papers has had a synergistic effect, commanding the attention of a broad swath of the biology community. "If a single paper comes out, it gets some attention, but when six papers come out all together, that's when people say, 'I have to do this,'" says Charles Gersbach, a biomedical engineer at Duke University in Durham, North Carolina.

Once she saw Doudna and Charpentier's paper a year ago, Gao Caixia became one of the early converts. Her group at the Chinese Academy of Sciences' Institute of Genetics and Developmental Biology in Beijing had been using zinc finger and TALENs technology on rice and wheat. Using CRISPR, they have now disabled four rice genes, suggesting that the technique could be used to engineer this crucial food crop. In wheat, they knocked out a gene that, when disabled, may lead to plants resistant to powdery mildew. In a measure of the excitement that CRISPR has generated, the team's report in the August 2013 issue of *Nature Biotechnology* was accompanied by four other papers describing CRISPR successes in plants and in rats.



CRISPRed rice. Earlier this month, researchers showed CRISPR works in plants, such as rice, where the knocked-out gene resulted in dwarf albino individuals (*right*).

PHOTO: GAO CAIXIA LABORATORY

The cost of admission is low: Free software exists to design guide RNA to target any desired gene, and a repository called Addgene, based in Cambridge, offers academics the DNA to make their own CRISPR system for $65. Since the beginning of the year, Addgene—to which 11 teams have contributed CRISPR-enabling DNA sequences—has distributed 5000 CRISPR constructs, and in a single July week the repository received 100 orders for a new construct. "They are kind of crazy hot," says Joanne Kamens, Addgene's executive director.

Fine-tuning gene activity

The initial CRISPR genome-editing papers all relied on DNA cutting, but other applications quickly appeared. Working with Doudna, Lei S. Qi from UC San Francisco and his colleagues introduced "CRISPRi," which, like RNAi, turns off genes in a reversible fashion and should be useful for studies of gene function. They modified Cas9 so it and the associated guide RNA would still home in on a target but would not cut DNA once there. In bacteria, the presence of Cas9 alone is enough to block transcription, but for mammalian applications, Qi and colleagues add to it a section of protein that represses gene activity. Its guide RNA is designed to home in on regulatory DNA, called promoters, which immediately precede the gene target.

Last month, that team and three other groups used a Cas9 to ferry a synthetic transcription factor—a protein fragment that turns on genes—enabling them to activate specific human genes. Just using one CRISPR construct had a weak effect, but all four teams found a way to amplify it. By targeting multiple CRISPR constructs to slightly different spots on the gene's promoter, says Gersbach, one of the team leaders, "we saw a huge synergistic effect."

In the 25 July 2013 issue of *Nature Methods*, he reported activating genes tied to human diseases, including those involved in muscle differentiation, controlling cancer and inflammation, and producing fetal hemoglobin. Two other teams also targeted biomedically important genes. CRISPR control of such genes could treat diseases ranging from sickle cell anemia to arthritis, Gersbach suggests.

CRISPR technology may yet have limitations. It's unclear, for example, how specific the guide RNAs are for just the genes they are supposed to target. "Our initial data suggest that there can be significant off-target effects," says J. Keith Joung from the Massachusetts General Hospital in Boston, who back in January demonstrated that CRISPR would alter genes in zebrafish embryos and has used CRISPR to turn on genes. His work shows that nontarget DNA resembling the guide RNA can become cut, activated, or deactivated.

Joung's group showed that a guide RNA can target DNA that differs from the intended target sequence in up to five of its bases. Zhang has gotten more reassuring results but says that "the specificity is still something we have to work on," especially as more people begin to think about delivering CRISPR systems as treatments for human diseases. "To really make the technology safe, we really have to make sure it goes where we want it to go to and nowhere else."

Researchers must also get the CRISPR components to the right place. "Delivery is an enormous challenge and will be cell type and organism specific," Joung notes. With zebrafish, his team injects guide RNA and messenger RNA for Cas9 directly into embryos; with mammalian cells, they use DNA constructs. How CRISPR might be delivered into adult animals, or to treat disease in people, is just now being considered.

Ultimately, CRISPR may take a place beside zinc fingers and TALENs, with the choice of editing tool depending on the particular application. But for now, researchers are dazzled by the ease by which they can make and test different CRISPR variants and by the technology's unexplored potential. Charpentier and others are looking at the versions of Cas9 in other bacteria that might work better than the one now being used. Microbiologists have harnessed the CRISPR system to vaccinate bacteria against the spread of antibiotic resistance genes. Church, Doudna, Charpentier, and others are forming CRISPR-related companies to begin exploring human therapeutic applications, including gene therapy.

And there's more that can be done, Barrangou says. "The only limitation today is people's ability to think of creative ways to harness [CRISPR]."

Not bad for a system that started with sickly bacteria.

Multiplex genome engineering using CRISPR/Cas systems

Le Cong,[1,2*] F. Ann Ran,[1,4,*] David Cox,[1,3] Shuailiang Lin,[1,5] Robert Barretto,[6] Naomi Habib,[1] Patrick D. Hsu,[1,4] Xuebing Wu,[7] Wenyan Jiang,[8] Luciano A. Marraffini,[8] Feng Zhang[1†]

Functional elucidation of causal genetic variants and elements requires precise genome editing technologies. The type II prokaryotic CRISPR (clustered regularly interspaced short palindromic repeats)/Cas adaptive immune system has been shown to facilitate RNA-guided site-specific DNA cleavage. We engineered two different type II CRISPR/Cas systems and demonstrate that Cas9 nucleases can be directed by short RNAs to induce precise cleavage at endogenous genomic loci in human and mouse cells. Cas9 can also be converted into a nicking enzyme to facilitate homology-directed repair with minimal mutagenic activity. Lastly, multiple guide sequences can be encoded into a single CRISPR array to enable simultaneous editing of several sites within the mammalian genome, demonstrating easy programmability and wide applicability of the RNA-guided nuclease technology.

Precise and efficient genome-targeting technologies are needed to enable systematic reverse engineering of causal genetic variations by allowing selective perturbation of individual genetic elements. Although genome-editing technologies such as designer zinc fingers (ZFs) (*1*–*4*), transcription activator–like effectors (TALEs) (*4*–*10*), and homing meganucleases (*11*) have begun to enable targeted genome modifications, there remains a need for new technologies that are scalable, affordable, and easy to engineer. Here, we report the development of a class of precision genome-engineering tools based on the RNA-guided Cas9 nuclease (*12*–*14*) from the type II prokaryotic clustered regularly interspaced short palindromic repeats (CRISPR) adaptive immune system (*15*–*18*).

The *Streptococcus pyogenes* SF370 type II CRISPR locus consists of four genes, including the Cas9 nuclease, as well as two noncoding CRISPR RNAs (crRNAs): trans-activating crRNA (tracrRNA) and a precursor crRNA (pre-crRNA) array containing nuclease guide sequences (spacers) interspaced by identical direct repeats (DRs) (fig. S1) (*19*). We sought to harness this prokaryotic RNA-programmable nuclease system to introduce targeted double-stranded breaks (DSBs) in mammalian chromosomes through heterologous expression of the key components. It has been previously shown that expression of tracrRNA, pre-crRNA, host factor ribonuclease (RNase) III, and Cas9 nuclease is necessary and sufficient for cleavage of DNA in vitro (*12*, *13*) and in prokaryotic cells (*20*, *21*). We codon-optimized the *S. pyogenes Cas9* (*SpCas9*) and *RNase III* (*SpRNase III*) genes and attached nuclear localization signals (NLSs) to ensure nuclear compartmentalization in mammalian cells. Expression of these constructs in human 293FT cells revealed that two NLSs are most efficient at targeting SpCas9 to the nucleus (Fig. 1A). To reconstitute the noncoding RNA components of the *S. pyogenes* type II CRISPR/Cas system, we expressed an 89-nucleotide (nt) tracrRNA (fig. S2) under the RNA polymerase III U6 promoter (Fig. 1B). Similarly, we used the U6 promoter to drive the expression of a pre-crRNA array comprising a

[1]Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA, and McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, Department of Biological Engineering, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. [2]Program in Biological and Biomedical Sciences, Harvard Medical School, Boston, MA 02115, USA. [3]Harvard-MIT Health Sciences and Technology, Harvard Medical School, Boston, MA 02115, USA. [4]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA. [5]School of Life Sciences, Tsinghua University, Beijing 100084, China. [6]Department of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA. [7]Computational and Systems Biology Graduate Program and Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [8]Laboratory of Bacteriology, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA.
*These authors contributed equally to this work.
†Corresponding author. E-mail: zhang@broadinstitute.org

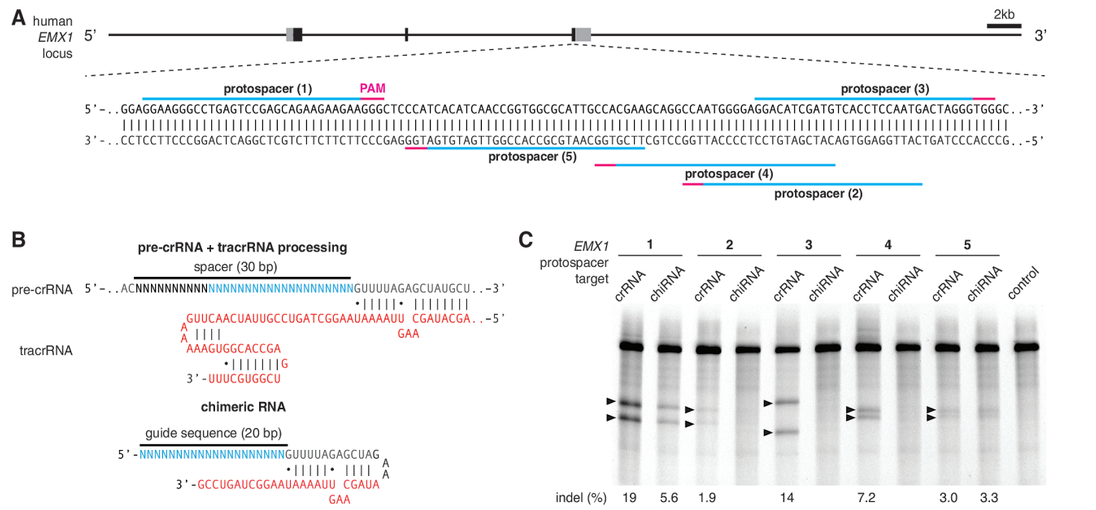


Fig. 2. SpCas9 can be reprogrammed to target multiple genomic loci in mammalian cells. (**A**) Schematic of the human *EMX1* locus showing the location of five protospacers indicated by blue lines with corresponding PAM in magenta. (**B**) Schematic of the pre-crRNA:tracrRNA complex (top) showing hybridization between the direct repeat (gray) region of the pre-crRNA and tracrRNA. Schematic of a chimeric RNA design (*12*) (bottom). tracrRNA sequence is shown in red and the 20-bp spacer sequence in blue. (**C**) SURVEYOR assay comparing the efficacy of Cas9-mediated cleavage at five protospacers in the human *EMX1* locus. Each protospacer was targeted by using either processed pre-crRNA:tracrRNA complex (crRNA) or chimeric RNA (chiRNA). Arrowheads indicate cleavage products for each protospacer target.

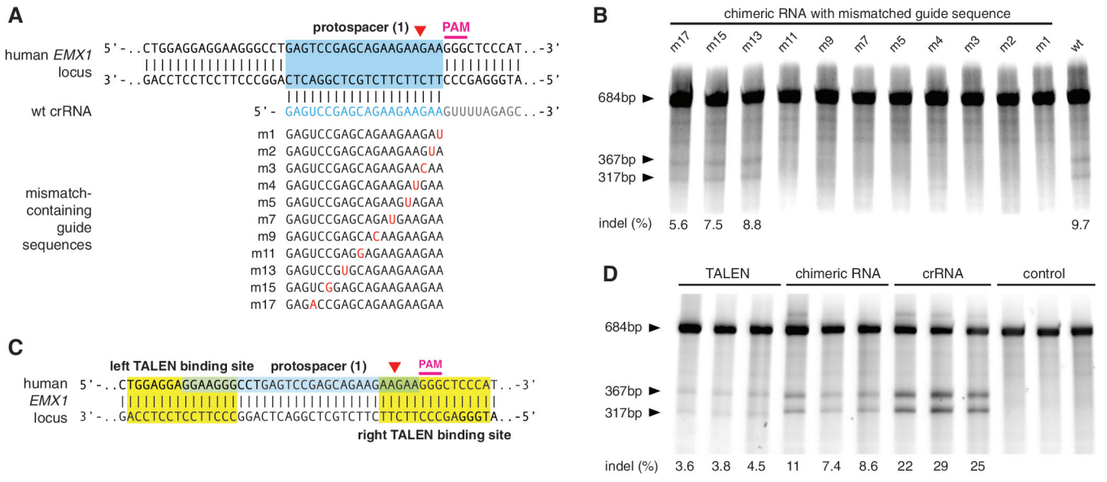


Fig. 3. Evaluation of the SpCas9 specificity and comparison of efficiency with TALENs. (**A**) *EMX1*-targeting chimeric crRNAs with single point mutations were generated to evaluate the effects of spacer-protospacer mismatches. (**B**) SURVEYOR assay comparing the cleavage efficiency of different mutant chimeric RNAs. (**C**) Schematic showing the design of TALENs that target *EMX1*. (**D**) SURVEYOR gel comparing the efficiency of TALEN andSpCas9 ($N = 3$).

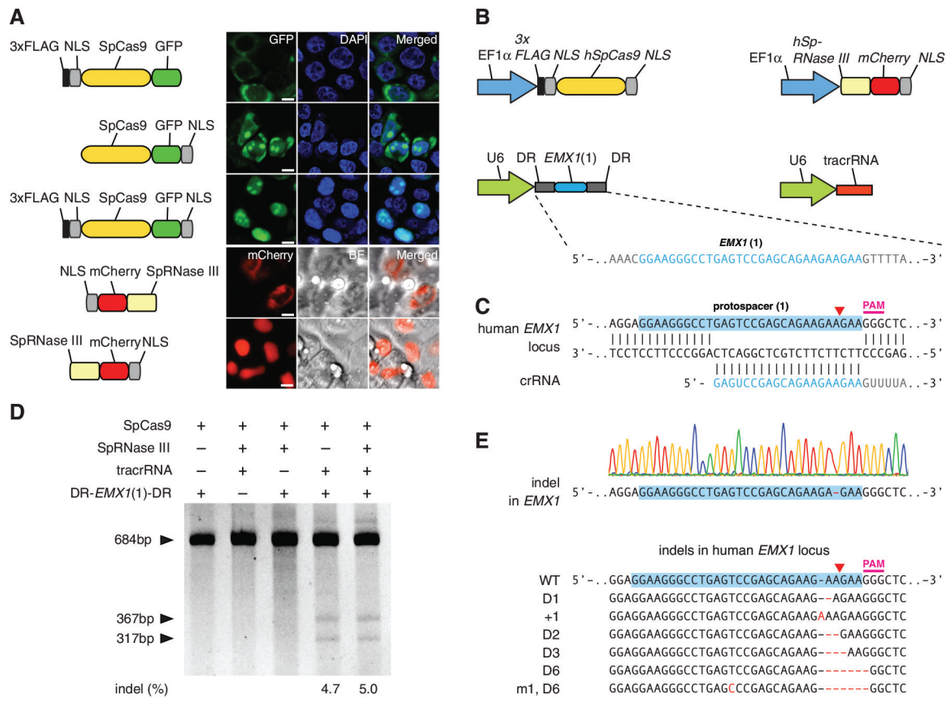


Fig. 1. The type II CRISPR locus from *S. pyogenes* SF370 can be reconstituted in mammalian cells to facilitate targeted DSBs of DNA. (**A**) Engineering of SpCas9 and SpRNase III with NLSs enables import into the mammalian nucleus. GFP indicates green fluorescent protein; scale bars, 10 μm. (**B**) Mammalian expression of human codon–optimized SpCas9 (hSpCas9) and SpRNase III (hSpRNase III) genes were driven by the elongation factor 1α (EF1α) promoter, whereas tracrRNA and pre-crRNA array (DR-Spacer-DR) were driven by the U6 promoter. A protospacer (blue highlight) from the human *EMX1* locus with PAM was used as template for the spacer in the pre-crRNA array. (**C**) Schematic representation of base pairing between target locus and *EMX1*-targeting crRNA. Red arrow indicates putative cleavage site. (**D**) SURVEYOR assay for SpCas9-mediated indels. (**E**) An example chromatogram showing a microdeletion, as well as representative sequences of mutated alleles identified from 187 clonal amplicons. Red dashes, deleted bases; red bases, insertions or mutations.

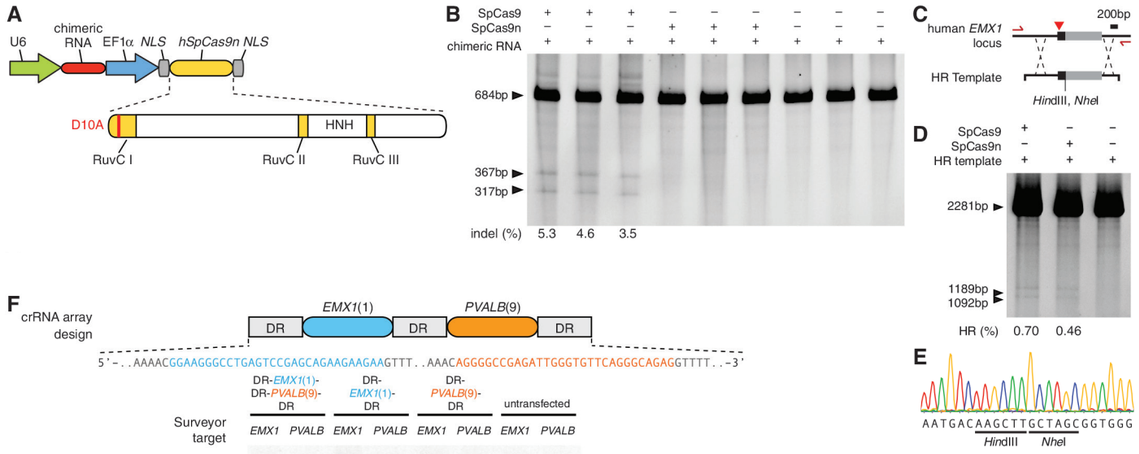


Fig. 4. Applications of Cas9 for homologous recombination and multiplex genome engineering. (**A**) Mutation of the RuvC I domain converts Cas9 into a nicking enzyme (SpCas9n). HNH, histidine-asparagine-histidine endonuclease domain. (**B**) Coexpression of *EMX1*-targeting chimeric RNA with SpCas9 leads to indels, whereas SpCas9n does not (N = 3). (**C**) Schematic representation of the recombination strategy. A homology repair (HR) template is designed to insert restriction sites into *EMX1* locus. Primers used to amplify the modified region are shown as red arrows. (**D**) Restriction fragment length polymorphism gel analysis. Arrows indicate fragments generated by *Hind*III digestion. (**E**) Example chromatogram showing successful recombination. (**F**) SpCas9 can facilitate multiplex genome modification by using a crRNA array that contains two spacers targeting *EMX1* and *PVALB*. Schematic showing the design of the crRNA array (top). Both spacers mediate efficient protospacer cleavage (bottom). (**G**) SpCas9 can be used to achieve precise genomic deletion. Two spacers targeting *EMX1* (top) mediated a 118-bp genomic deletion (bottom).

single guide spacer flanked by DRs (Fig. 1B). We designed our initial spacer to target a 30–base pair (bp) site (protospacer) in the human *EMX1* locus that precedes an NGG trinucleotide, the requisite protospacer-adjacent motif (PAM) (Fig. 1C and fig. S1) (*22*, *23*).

To test whether heterologous expression of the CRISPR system (SpCas9, SpRNase III, tracrRNA, and pre-crRNA) can achieve targeted cleavage of mammalian chromosomes, we transfected 293FT cells with different combinations of CRISPR/Cas components. Because DSBs in mammalian DNA are partially repaired by the indel-forming nonhomologous end joining (NHEJ) pathway, we used the SURVEYOR assay (fig. S3) to detect endogenous target cleavage (Fig. 1D and fig. S2B). Cotransfection of all four required CRISPR components resulted in efficient cleavage of the protospacer (Fig. 1D and fig. S2B), which was subsequently verified by Sanger sequencing (Fig. 1E). SpRNase III was not necessary for cleavage of the protospacer (Fig. 1D), and the 89-nt tracrRNA is processed in its absence (fig. S2C). Similarly, maturation of pre-crRNA does not require RNase III (Fig. 1D and fig. S4), suggesting that there may be endogenous mammalian RNases that assist in pre-crRNA maturation (*24*–*26*). Removing any of the remaining RNA or Cas9 components abolished the genome cleavage activity of the CRISPR/Cas system (Fig. 1D). These results define a minimal three-component system for efficient RNA-guided genome modification in mammalian cells.

Next, we explored the generalizability of RNA-guided genome editing in eukaryotic cells by targeting additional protospacers within the *EMX1* locus (Fig. 2A). To improve codelivery, we designed an expression vector to drive both pre-crRNA and SpCas9 (fig. S5). In parallel, we adapted a chimeric crRNA-tracrRNA hybrid (Fig. 2B, top) design recently validated in vitro (*12*), where a mature crRNA is fused to a partial tracrRNA via a synthetic stem loop to mimic the natural crRNA:tracrRNA duplex (Fig. 2B, bottom). We observed cleavage of all protospacer targets when SpCas9 is coexpressed with pre-crRNA (DR-spacer-DR) and tracrRNA. However, not all chimeric RNA designs could facilitate cleavage of their genomic targets (Fig. 2C and table S1). We then tested targeting of additional genomic loci in both human and mouse cells by designing pre-crRNAs and chimeric RNAs targeting the human *PVALB* and the mouse *Th* loci (fig. S6). We achieved efficient modification at all three mouse *Th* and one *PVALB* targets by using the crRNA:tracrRNA duplex, thus demonstrating the broad applicability of the CRISPR/Cas system in modifying different loci across multiple organisms (table S1). For the same protospacer targets, cleavage efficiencies of chimeric RNAs were either lower than those of crRNA:tracrRNA duplexes or undetectable. This may be due to differences in the expression and stability of RNAs, degradation by endogenous RNA interference machinery, or secondary structures leading to inefficient Cas9 loading or target recognition.

Effective genome editing requires that nucleases target specific genomic loci with both high precision and efficiency. To investigate the specificity of RNA-guided genome modification, we analyzed single-nucleotide mismatches between the spacer and its mammalian protospacer target (Fig. 3A). We observed that single-base mismatch up to 11 bp 5′ of the PAM completely abolished genomic cleavage by SpCas9, whereas spacers with mutations farther upstream retained activity against the protospacer target (Fig. 3B). This is consistent with previous bacterial and in vitro studies of Cas9 specificity (*12*, *20*). Furthermore, SpCas9 is able to mediate genomic cleavage as efficiently as a pair of TALE nucleases (TALENs) targeting the same *EMX1* protospacer (Fig. 3, C and D).

Targeted modification of genomes ideally avoids mutations arising from the error-prone NHEJ mechanism. The wild-type SpCas9 is able to mediate site-specific DSBs, which can be repaired through either NHEJ or homology-directed repair (HDR). We engineered an aspartate-to-alanine substitution (D10A) in the RuvC I domain of SpCas9 to convert the nuclease into a DNA nickase (SpCas9n, Fig. 4A) (*12*, *13*, *20*), because nicked genomic DNA is typically repaired either seamlessly or through high-fidelity HDR. SURVEYOR (Fig. 4B) and sequencing of 327 amplicons did not detect any indels induced by SpCas9n. However, nicked DNA can in rare cases be processed via a DSB intermediate and result in a NHEJ event (*27*). We then tested Cas9-mediated HDR at the same *EMX1* locus with a homology repair template to introduce a pair of restriction sites near the protospacer (Fig. 4C). SpCas9 and SpCas9n catalyzed integration of the repair template into *EMX1* locus at similar levels (Fig. 4D), which we further verified via Sanger sequencing (Fig. 4E). These results demonstrate the utility of CRISPR for facilitating targeted genomic insertions. Given the 14-bp (12 bp from the seed sequence and 2 bp from PAM) target specificity (Fig. 3B) of the wild-type SpCas9, the use of a nickase may reduce off-target mutations.

Lastly, the natural architecture of CRISPR loci with arrayed spacers (fig. S1) suggests the possibility of multiplexed genome engineering. By using a single CRISPR array encoding a pair of *EMX1*- and *PVALB*-targeting spacers, we detected efficient cleavage at both loci (Fig. 4F). We further tested targeted deletion of larger genomic regions through concurrent DSBs by using spacers against two targets within *EMX1* spaced by 119 bp and observed a 1.6% deletion efficacy (3 out of 182 amplicons, Fig. 4G), thus demonstrating the CRISPR/Cas system can mediate multiplexed editing within a single genome.

The ability to use RNA to program sequence-specific DNA cleavage defines a new class of genome engineering tools. Here, we have shown that the *S. pyogenes* CRISPR system can be heterologously reconstituted in mammalian cells to facilitate efficient genome editing; an accompanying study has independently confirmed high-efficiency RNA-guided genome targeting in several human cell lines (*28*). However, several aspects of the CRISPR/Cas system can be further improved to increase its efficiency and versatility. The requirement for an NGG PAM restricts the target space of SpCas9 to every 8 bp on average in the human genome (fig. S7), not accounting for potential constraints posed by crRNA secondary structure or genomic accessibility resulting from chromatin and DNA methylation states. Some of these restrictions may be overcome by exploiting the family of Cas9 enzymes and its differing PAM requirements (*22*, *23*) across the microbial diversity (*17*). Indeed, other CRISPR loci are likely to be transplantable into mammalian cells; for example, the *Streptococcus thermophilus* LMD-9 CRISPR1 system can also mediate mammalian genome cleavage (fig. S8). Lastly, the ability to carry out multiplex genome editing in mammalian cells enables powerful applications across basic science, biotechnology, and medicine (*29*).

REFERENCES AND NOTES

1. M. H. Porteus, D. Baltimore, *Science* **300**, 763 (2003).
2. J. C. Miller *et al.*, *Nat. Biotechnol.* **25**, 778 (2007).
3. J. D. Sander *et al.*, *Nat. Methods* **8**, 67 (2011).
4. A. J. Wood *et al.*, *Science* **333**, 307 (2011).
5. M. Christian *et al.*, *Genetics* **186**, 757 (2010).
6. F. Zhang *et al.*, *Nat. Biotechnol.* **29**, 149 (2011).
7. J. C. Miller *et al.*, *Nat. Biotechnol.* **29**, 143 (2011).
8. D. Reyon *et al.*, *Nat. Biotechnol.* **30**, 460 (2012).
9. J. Boch *et al.*, *Science* **326**, 1509 (2009).
10. M. J. Moscou, A. J. Bogdanove, *Science* **326**, 1501 (2009).
11. B. L. Stoddard, Q. *Rev. Biophys.* **38**, 49 (2005).
12. M. Jinek *et al.*, *Science* **337**, 816 (2012).
13. G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, *Proc. Natl. Acad. Sci. U.S.A.* **109**, E2579 (2012).
14. J. E. Garneau *et al.*, *Nature* **468**, 67 (2010).
15. H. Deveau, J. E. Garneau, S. Moineau, *Annu. Rev. Microbiol.* **64**, 475 (2010).
16. P. Horvath, R. Barrangou, *Science* **327**, 167 (2010).
17. K. S. Makarova *et al.*, *Nat. Rev. Microbiol.* **9**, 467 (2011).
18. D. Bhaya, M. Davison, R. Barrangou, *Annu. Rev. Genet.* **45**, 273 (2011).
19. E. Deltcheva *et al.*, *Nature* **471**, 602 (2011).
20. R. Sapranauskas *et al.*, *Nucleic Acids Res.* **39**, 9275 (2011)
21. A. H. Magadán, M. E. Dupuis, M. Villion, S. Moineau, *PLOS ONE* **7**, e40913 (2012).
22. H. Deveau *et al.*, *J. Bacteriol.* **190**, 1390 (2008).
23. F. J. Mojica, C. Díez-Villaseñor, J. García-Martínez, C. Almendros, *Microbiology* **155**, 733 (2009).
24. M. Jinek, J. A. Doudna, *Nature* **457**, 405 (2009).
25. C. D. Malone, G. J. Hannon, *Cell* **136**, 656 (2009).
26. G. Meister, T. Tuschl, *Nature* **431**, 343 (2004).
27. M. T. Certo *et al.*, *Nat. Methods* **8**, 671 (2011).
28. P. Mali *et al.*, *Science* **339**, 823 (2013).
29. P. A. Carr, G. M. Church, *Nat. Biotechnol.* **27**, 1151 (2009).

ACKNOWLEDGMENTS

We thank the entire Zhang lab for their support and advice; P. A. Sharp for generous help with Northern blot analysis; C. Jennings, R. Desimone, and M. Kowalczyk for helpful comments; and X. Ye for help with confocal imaging. L.C. and X.W. are Howard Hughes Medical Institute International Student Research Fellows. D.C. is supported by the Medical Scientist Training Program. P.D.H. is a James Mills Pierce Fellow. X.W. is supported by NIH grants R01-GM34277 and R01-CA133404 to P. A. Sharp. X.W.'s thesis adviser. L. A. M. is supported by Searle Scholars, R. Allen, an Irma T. Hirschl Award, and a NIH Director's New Innovator Award (DP2AI104556). F.Z. is supported by a NIH Director's Pioneer Award (DP1MH100706); the Keck, McKnight, Gates, Damon Runyon, Searle Scholars, Klingenstein, and Simons foundations; R. Metcalfe; M. Boylan; and J. Pauley. The authors have no conflicting financial interests. A patent application has been filed relating to this work, and the authors plan on making the reagents widely available to the academic community through Addgene and to provide software tools via the Zhang lab Web site (www.genome-engineering.org).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/cgi/content/full/science.1231143/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 and S2
References (*30*–*32*)

5 October 2012; accepted 12 December 2012
Published online 3 January 2013;
10.1126/science.1231143

RNA-guided human genome engineering via Cas9

Prashant Mali,[1*] Luhan Yang,[1,3*] Kevin M. Esvelt,[2] John Aach,[1] Marc Guell,[1] James E. DiCarlo,[4] Julie E. Norville,[1] George M. Church,[1,2†]

Bacteria and archaea have evolved adaptive immune defenses, termed clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems, that use short RNA to direct degradation of foreign nucleic acids. Here, we engineer the type II bacterial CRISPR system to function with custom guide RNA (gRNA) in human cells. For the endogenous AAVS1 locus, we obtained targeting rates of 10 to 25% in 293T cells, 13 to 8% in K562 cells, and 2 to 4% in induced pluripotent stem cells. We show that this process relies on CRISPR components; is sequence-specific; and, upon simultaneous introduction of multiple gRNAs, can effect multiplex editing of target loci. We also compute a genome-wide resource of ~190 K unique gRNAs targeting ~40.5% of human exons. Our results establish an RNA-guided editing tool for facile, robust, and multiplexable human genome engineering.

Bacterial and archaeal clustered regularly interspaced short palindromic repeats (CRISPR) systems rely on CRISPR RNAs (crRNAs) in complex with CRISPR-associated (Cas) proteins to direct degradation of complementary sequences present within invading viral and plasmid DNA (*1–3*). A recent in vitro reconstitution of the *Streptococcus pyogenes* type II CRISPR system demonstrated that crRNA fused to a normally trans-encoded tracrRNA is sufficient to direct Cas9 protein to sequence-specifically cleave target DNA sequences matching the crRNA (*4*). The fully defined nature of this two-component system suggested that it might function in the cells of eukaryotic organisms such as yeast, plants, and even mammals. By cleaving genomic sequences targeted by RNA sequences (*4–6*), such a system could greatly enhance the ease of genome engineering.

Here, we engineer the protein and RNA components of this bacterial type II CRISPR system in human cells. We began by synthesizing a human codon–optimized version of the Cas9 protein bearing a C-terminal SV40 nuclear localization signal and cloning it into a mammalian expression system (Fig. 1A and fig. S1A). To direct Cas9 to cleave sequences of interest, we expressed crRNA-tracrRNA fusion transcripts, hereafter referred to as guide RNAs (gRNAs), from the human U6 polymerase III promoter. Directly transcribing gRNAs allowed us to avoid reconstituting the RNA-processing machinery used by bacterial CRISPR systems (Fig. 1A and fig. S1B) (*4, 7–9*). Constrained only by U6 transcription initiating with G and the requirement for the PAM (protospacer-adjacent motif) sequence -NGG following the 20–base pair (bp) crRNA target, our highly versatile approach can, in principle, target any genomic site of the form $GN_{20}GG$ (fig. S1C; see supplementary text S1 for a detailed discussion).

To test the functionality of our implementation for genome engineering, we developed a green fluorescent protein (GFP) reporter assay (Fig. 1B) in human embryonic kidney HEK 293T cells similar to one previously described (*10*). Specifically, we established a stable cell line bearing a genomically integrated GFP coding sequence disrupted by the insertion of a stop codon and a 68-bp genomic fragment from the AAVS1 locus that renders the expressed protein fragment nonfluorescent. Homologous recombination (HR) using an appropriate repair donor can restore the normal GFP sequence, which enabled us to quantify the resulting GFP^+ cells by flow-activated cell sorting (FACS).

To test the efficiency of our system at stimulating HR, we constructed two gRNAs, T1 and T2, that target the intervening AAVS1 fragment (Fig. 1B) and compared their activity to that of a previously described TAL effector nuclease heterodimer (TALEN) targeting the same region (*11*). We observed successful HR events using all three targeting reagents, with gene correction rates using the T1 and T2 gRNAs approaching 3% and 8%, respectively (Fig. 1C). This RNA-mediated editing process was notably rapid, with the first detectable GFP^+ cells appearing ~20 hours post transfection compared with ~40 hours for the AAVS1 TALENs. We observed HR only upon simultaneous introduction of the repair donor, Cas9 protein, and gRNA, which confirmed that all components are required for genome editing (fig. S2). Although we noted no apparent toxicity associated with Cas9/gRNA expression, work with zinc finger nucleases (ZFNs) and TALENs has shown that nicking only one strand further reduces toxicity. Accordingly, we also tested a Cas9D10A mutant that is known to function as a nickase in vitro, which yielded similar HR but lower nonhomologous end joining (NHEJ) rates (fig. S3) (*4, 5*). Consistent with (*4*), in which a related Cas9 protein is shown to cut both strands 3 bp upstream of the PAM, our NHEJ data confirmed that most deletions or insertions occurred at the 3′ end of the target sequence (fig. S3B). We also confirmed that mutating the target genomic site prevents the gRNA from effecting HR at that locus, which demonstrates that CRISPR-mediated genome editing is sequence-specific (fig. S4). Finally, we showed that two gRNAs targeting sites in the GFP gene, and also three additional gRNAs targeting fragments from homologous regions of the DNA methyl transferase 3a (DNMT3a) and DNMT3b genes could sequence-specifically induce significant HR in the engineered reporter cell lines (figs. S5 and S6). Together, these results confirm that RNA-guided genome targeting in human cells is simple to execute and induces robust HR across multiple target sites.

Having successfully targeted an integrated reporter, we next turned to modifying a native locus. We used the gRNAs described above to target the AAVS1 locus located in the PPP1R12C gene on chromosome 19, which is ubiquitously expressed across most tissues (Fig. 2A). We targeted 293Ts, human chronic myelogenous leukemia K562 cells, and PGP1 human induced pluripotent stem (iPS) cells (*12*) and analyzed the results by next-generation sequencing of the targeted locus. Consistent with our results for the GFP reporter assay, we observed high numbers of NHEJ events at the endogenous locus for all three cell types. The two gRNAs T1 and T2 achieved NHEJ rates of 10 and 25% in 293Ts, 13 and 38% in K562s, and 2 and 4% in PGP1-iPS cells, respectively (Fig. 2B). We observed no overt toxicity from the Cas9 and gRNA expression required to induce NHEJ in any of these cell types. As expected, NHEJ-mediated deletions for T1 and T2 were centered around the target site positions, which further validated the sequence-specificity of this targeting process (figs. S7 to S9). Simultaneous introduction of both T1 and T2 gRNAs resulted in high-efficiency deletion of the intervening 19-bp fragment (fig. S8), which demonstrated that multiplexed editing of genomic loci is feasible using this approach.

Last, we attempted to use HR to integrate either a double-stranded DNA donor construct (*13*) or an oligo donor into the native AAVS1 locus (Fig. 2C and fig. S10). We confirmed HR-mediated integration, using both approaches, by polymerase chain reaction (PCR) (Fig. 2D and fig. S10) and Sanger sequencing (Fig. 2E). We also readily derived 293T or iPS clones from the pool of modified cells using puromycin selection over 2 weeks (Fig. 2F and fig. S10). These results demonstrate that this approach enables efficient integration of foreign DNA at endogenous loci in human cells.

Our versatile RNA-guided genome-editing system can be readily adapted to modify other genomic sites by simply modifying the sequence of our gRNA expression vector to match a compatible sequence in the locus of interest. To facilitate this process, we bioinformatically generated ~190,000 specific gRNA-targetable sequences targeting ~40.5% exons of genes in the human genome (refer to methods and table S1). We also incorporated these target sequences into a 200-bp format compatible with multiplex synthesis on DNA arrays (*14*) (fig. S11 and tables S2 and S3). This resource provides a ready genome-wide reference of potential target sites in the human genome and a methodology for multiplex gRNA synthesis.

Our results demonstrate the promise of CRISPR-mediated gene targeting for RNA-guided, robust, and multiplexable mammalian genome engineering. The ease of retargeting our system to modify genomic sequences greatly exceeds that of comparable ZFNs and TALENs, while offering similar or greater efficiencies (*4*). Existing studies of type II CRISPR specificity (*4*) suggest that target sites must perfectly match the PAM sequence NGG and the 8- to 12-base "seed sequence" at the 3′ end of the gRNA. The importance of the remaining 8 to 12 bases is less well understood and may depend on the binding strength of the matching gRNAs or on the inherent tolerance of Cas9 itself. Indeed, Cas9 will tolerate single mismatches at the 5′ end in bacteria and in vitro, which suggests that the 5′ G is not required. Moreover, it is likely that the target locus's underlying chromatin structure and epigenetic state will also affect the efficiency of genome editing in eukaryotic cells (*13*), although we suspect that Cas9's helicase activity may render it more robust to these factors, but this remains to be evaluated. Elucidating the frequency and underlying causes of off-target nuclease activity (*15, 16*) induced by CRISPR, ZFN (*17, 18*), and TALEN (*19, 20*) genome-engineering tools will be of utmost importance for safe genome modification and perhaps for gene therapy. Potential avenues for improving CRISPR specificity include evaluating Cas9 homologs identified through bioinformatics and directed evolution of these nucleases toward higher specificity. Similarly, the range of CRISPR-targetable sequences could be expanded through the use of homologs with different PAM requirements (*9*) or by directed evolution. Finally, inactivating one of the Cas9 nuclease domains increases the ratio of HR to NHEJ and may reduce toxicity (figs. S1A and fig. S3) (*4, 5*), whereas inactivating both domains may enable Cas9 to function as a retargetable DNA binding protein. As we explore these areas, we note that another parallel study (*21*) has independently confirmed the high efficiency of CRISPR-mediated gene targeting in mammalian cell lines. We expect that RNA-guided genome targeting will have broad implications for synthetic biology (*22, 23*), the direct and multiplexed perturbation of gene networks (*13, 24*), and targeted ex vivo (*25–27*) and in vivo gene therapy (*28*).

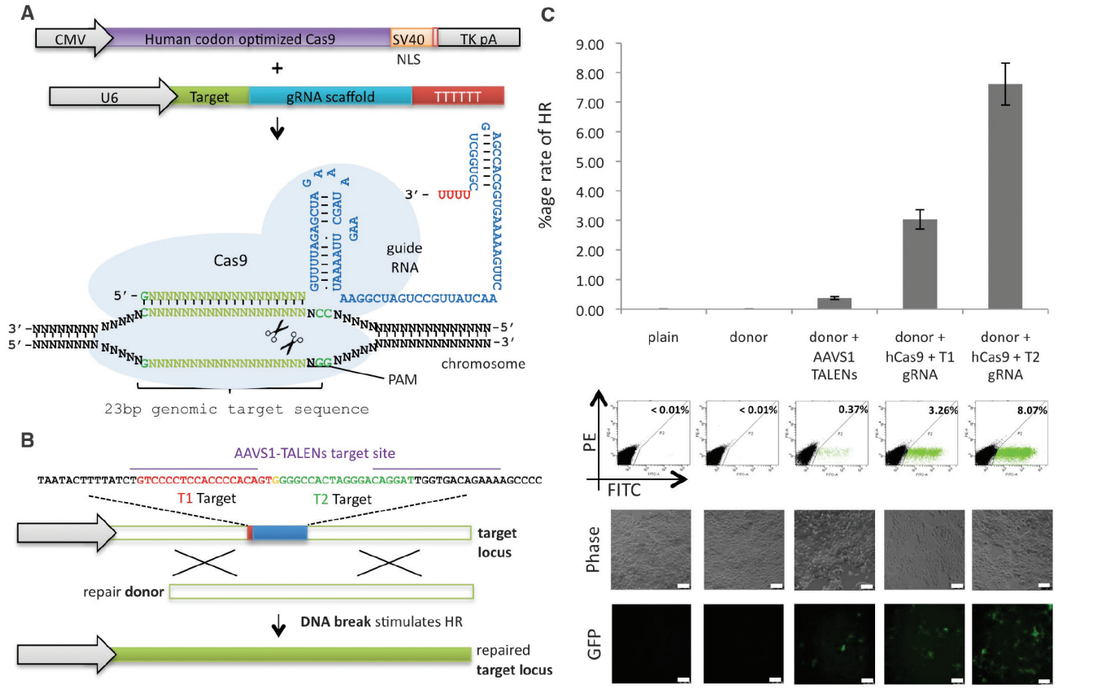


Fig. 1. Genome editing in human cells using an engineered type II CRISPR system. (**A**) RNA-guided gene targeting in human cells involves coexpression of the Cas9 protein bearing a C-terminal SV40 nuclear localization signal (NLS) with one or more gRNAs expressed from the human U6 polymerase III promoter. Cas9 unwinds the DNA duplex and cleaves both strands upon recognition of a target sequence by the gRNA, but only if the correct PAM is present at the 3′ end. Any genomic sequence of the form $GN_{20}GG$ can, in principle, be targeted. CMV, cytomegalovirus promoter; TK, thymidine kinase; pA, polyadenylation signal. (**B**) A genomically integrated GFP coding sequence is disrupted by the insertion of a stop codon and a 68-bp genomic fragment from the AAVS1 locus. Restoration of the GFP sequence by HR with an appropriate donor sequence results in GFP^+ cells that can be quantified by FACS. T1 and T2 gRNAs target sequences within the AAVS1 fragment. Binding sites for the two halves of the TALEN are underlined. (**C**) Bar graph depicting HR efficiencies induced by T1, T2, and TALEN-mediated nuclease activity at the target locus, as measured by FACS. Representative FACS plots and microscopy images of the targeted cells are depicted below. (Scale bar, 100 μm.) Data are shown as means ± SEM ($N = 3$).

[1]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA.
[2]Wyss Institute for Biologically Inspired Engineering, Harvard University, Cambridge, MA 02138, USA.
[3]Biological and Biomedical Sciences Program, Harvard Medical School, Boston, MA 02115, USA.
[4]Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA.
*These authors contributed equally to this work.
†**Corresponding author. E-mail: gchurch@genetics.med.harvard.edu**

REFERENCES AND NOTES
1. B. Wiedenheft, S. H. Sternberg, J. A. Doudna, *Nature*

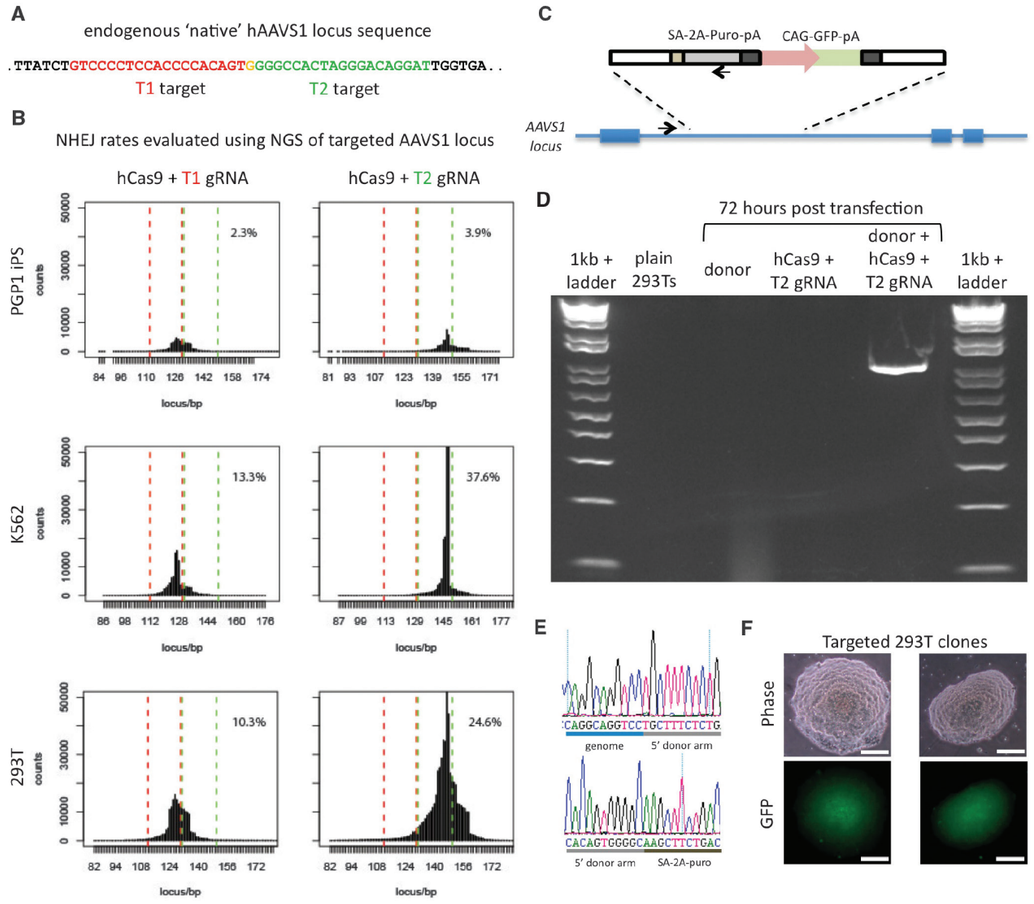


Fig. 2. RNA-guided genome editing of the native AAVS1 locus in multiple cell types. (**A**) T1 (red) and T2 (green) gRNAs target sequences in an intron of the PPP1R12C gene within the chromosome 19 AAVS1 locus. (**B**) Total count and location of deletions caused by NHEJ in 293Ts, K562s, and PGP1 iPS cells after expression of Cas9 and either T1 or T2 gRNAs as quantified by next-generation sequencing. Red and green dashed lines demarcate the boundaries of the T1 and T2 gRNA targeting sites. NHEJ frequencies for T1 and T2 gRNAs were 10% and 25% in 293T, 13% and 38% in K562, and 2% and 4% in PGP1 iPS cells, respectively. (**C**) DNA donor architecture for HR at the AAVS1 locus, and the locations of sequencing primers (arrows) for detecting successful targeted events, are depicted. (**D**) PCR assay 3 days after transfection demonstrates that only cells expressing the donor, Cas9 and T2 gRNA exhibit successful HR events. (**E**) Successful HR was confirmed by Sanger sequencing of the PCR amplicon, which showed that the expected DNA bases at both the genome-donor and donor-insert boundaries are present. (**F**) Successfully targeted clones of 293T cells were selected with puromycin for 2 weeks. Microscope images of two representative GFP+ clones is shown. (Scale bar, 100 μm.)

482, 331 (2012).
2. D. Bhaya, M. Davison, R. Barrangou, *Annu. Rev. Genet.* **45**, 273 (2011).
3. M. P. Terns, R. M. Terns, *Curr. Opin. Microbiol.* **14**, 321 (2011).
4. M. Jinek *et al.*, *Science* **337**, 816 (2012).
5. G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, *Proc. Natl. Acad. Sci. U.S.A.* **109**, E2579 (2012).
6. R. Sapranauskas *et al.*, *Nucleic Acids Res.* **39**, 9275 (2011).
7. T. R. Brummelkamp, R. Bernards, R. Agami, *Science* **296**, 550 (2002).
8. M. Miyagishi, K. Taira, *Nat. Biotechnol.* **20**, 497 (2002).
9. E. Deltcheva *et al.*, *Nature* **471**, 602 (2011).
10. J. Zou, P. Mali, X. Huang, S. N. Dowey, L. Cheng, *Blood* **118**, 4599 (2011).
11. N. E. Sanjana *et al.*, *Nat. Protoc.* **7**, 171 (2012).
12. J. H. Lee *et al.*, *PLoS Genet.* **5**, e1000718 (2009).
13. D. Hockemeyer *et al.*, *Nat. Biotechnol.* **27**, 851 (2009).
14. S. Kosuri *et al.*, *Nat. Biotechnol.* **28,** 1295 (2010).
15. V. Pattanayak, C. L. Ramirez, J. K. Joung, D. R. Liu, *Nat. Methods* **8**, 765 (2011).
16. N. M. King, O. Cohen-Haguenauer, *Mol. Ther.* **16**, 432 (2008).
17. Y. G. Kim, J. Cha, S. Chandrasegaran, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1156 (1996).
18. E. J. Rebar, C. O. Pabo, *Science* **263**, 671 (1994).
19. J. Boch *et al.*, *Science* **326**, 1509 (2009).
20. M. J. Moscou, A. J. Bogdanove, *Science* **326**, 1501 (2009).
21. L. Cong *et al.*, *Science* **339**, 819 (2013).
22. A. S. Khalil, J. J. Collins, *Nat. Rev. Genet.* **11**, 367 (2010).
23. P. E. Purnick, R. Weiss, *Nat. Rev. Mol. Cell Biol.* **10**, 410 (2009).
24. J. Zou *et al.*, *Cell Stem Cell* **5**, 97 (2009).
25. N. Holt *et al.*, *Nat. Biotechnol.* **28**, 839 (2010).
26. F. D. Urnov *et al.*, *Nature* **435**, 646 (2005).
27. A. Lombardo *et al.*, *Nat. Biotechnol.* **25**, 1298 (2007).
28. H. Li *et al.*, *Nature* **475**, 217 (2011).

ACKNOWLEDGMENTS

This work was supported by NIH grant P50 HG005550. We thank S. Kosuri for advice on the oligonucleotide pool designs and synthesis. G.M.C. and P.M.have filed a patent based on the findings of this study.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/cgi/content/full/science.1232033/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S11
Tables S1 to S3
References (*29*–*46*)

26 October 2012; accepted 12 December 2012
Published online 3 January 2013;
10.1126/science.1232033

A Swiss army knife of immunity

Stan J. J. Brouns

Selfish genetic elements are more than a daily nuisance in the life of prokaryotes. Whereas viruses can multiply by reprogramming host cells, or integrate in the host genome as "stowaways," conjugative plasmids (transferrable extrachromosomal DNA) make cells addicted to plasmid-encoded antitoxin factors, thus preventing their disposal. Bacteria and archaea defend themselves against these invasive elements using an adaptive immune system based on clustered regularly interspaced short palindromic repeats (CRISPRs). On page [30 in this booklet], Jinek *et al.* (*1*) show how the CRISPR effector enzyme Cas9 from bacteria is directed not by one, but two small RNAs to cleave invader DNA.

The CRISPR system integrates short DNA fragments from viruses and plasmids into a specific repeat locus of the host cell genome to function as a memory of past invasions. This locus of the "cell's most wanted" is then transcribed into RNA (the precursor CRISPR RNA), which is cleaved in each repeat to yield individual mature CRISPR RNAs (crRNAs). These guide a dedicated set of CRISPR-associated (Cas) proteins to their targets during cellular surveillance of the cytoplasm for either foreign DNA or messenger RNA (mRNA) of known invaders. Once identified, foreign nucleic acids are permanently damaged by Cas nucleases, thereby neutralizing the invader (*2*).

The CRISPR field was set in motion 5 years ago by the discovery that lactic acid bacteria become highly resistant to virus infection when they incorporate virus DNA fragments in their array of memorized invaders (*3*). Bacterial resistance to the virus is based on breaks in the viral DNA within this memorized region, and the bacterial gene *cas9* encodes the enzyme responsible (*4, 5*). However, the modus operandi of Cas9 has remained unknown.

One aspect that had to be resolved first was the unusual way in which Cas9 obtains the mature crRNA. Whereas most CRISPR-Cas systems involve a dedicated nuclease that cleaves the precursor CRISPR RNA in each repeat (*2*), Cas9-based systems also require a CRISPR-specific small RNA. This so-called trans-activated crRNA (tracrRNA) base pairs with each repeat of the CRISPR transcript and provides a substrate for the RNA-specific host ribonuclease RNase III (*6*). The cleavage product, an RNA hybrid consisting of a 42-nucleotide crRNA and a 75-nucleotide tracrRNA, was deemed to be the guide for Cas9.

With this in mind, Jinek *et al.* could show that Cas9 from the human pathogenic bacterium *Streptococcus pyogenes* binds and cleaves invader DNA within the remembered region. Although the site specificity was solely determined by the guiding ability of the crRNA, binding and cleavage of the target DNA surprisingly required the tracrRNA. The tracrRNA thus enables the Cas9-crRNA complex to locate a DNA sequence complementary to the crRNA in the cellular tangle of DNA (see the figure), providing yet another example of the crucial roles that small RNAs play in cells (*7*).

Cas9 creates blunt-ended lesions in target DNA by using two nuclease domains, each cleaving one DNA strand of the target double-stranded DNA R-loop. Whereas the HNH-nuclease domain cleaves the DNA strand that base pairs with the crRNA, the RuvC-nuclease domain cleaves the displaced strand of the DNA. Jinek *et al.* show that cleavage was robust and occurred with multiple turnovers in both relaxed and supercoiled DNA targets, implying that Cas9 is functionally recycled after cleavage to destroy more invader DNA copies that may be present in the host cell.

Despite the seeming efficiency of this cleavage and recycling process, the Achilles' heel of Cas9 was also uncovered. Viruses escape immunity by making point mutations in either the memorized regions of their genomes (*8*), or just outside this region in a conserved nucleotide motif. When testing these mutant DNA molecules, Jinek *et al.* found that binding and cleavage by Cas9 was severely compromised, suggesting that these mutated virus DNA molecules adopt a stealth mode inside the cell and require a new cycle of memory formation before they are subject to interference once again. Cycles like these contribute to the ongoing coevolution between invaders and their hosts.

Jinek *et al.* realized that a highly specific, customizable RNA-directed DNA nuclease could be useful to edit whole genomes. Based on the 20-nucleotide guide section of the crRNA, the enzyme could theoretically introduce breaks at unique sites in any eukaryotic genome. As a proof of concept, the authors programmed Cas9 to cleave a plasmid carrying the gene encoding green fluorescent protein at predetermined loci using a single chimeric crRNA containing just the critical segment of the tracrRNA. DNA breaks induce cellular DNA repair pathways (*9*) and this can be harnessed to disrupt, insert, or repair specific genes of cells. Introducing DNA breaks at desired loci using just Cas9 and a chimeric crRNA would be a substantial improvement over existing gene-targeting technologies, such as zinc finger nucleases and transcription activator–like effector nucleases, as these require protein engineering for every new target locus (*10*). Efficient gene repair strategies in cells from patients, and the reintroduction of repaired cells, could become increasingly important for treating many genetic disorders.

Cas9 is thus a remarkably compact and multifunctional enzyme compared to CRISPR effector complexes from other bacteria or archaea. These are typically 350- to 450-kD crRNA-protein complexes and contain up to 11 protein subunits encoded by four to seven different *cas* genes (see the figure) (*2*, *11*). Yet, nuclease activities are not always part of the complex and need to be recruited when the target DNA is identified (*12*). With all its activities at hand, Cas9

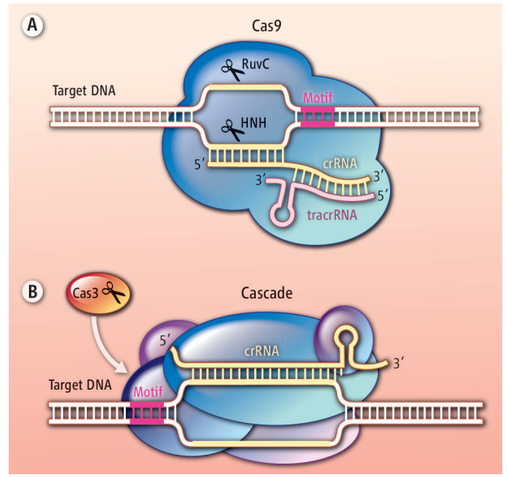


Fig. 1. All-in-one nuclease. (**A**) Cas9 requires a crRNA and tracrRNA to recognize invader DNA sequences by hybridizing the guide section of the crRNA to one strand of the target DNA to form an R-loop. The flanking motif is critical for this process and may facilitate DNA duplex unwinding and strand invasion by the crRNA. Target DNA is then cleaved by both nuclease domains of Cas9. (**B**) Cascade-like complexes contain a single crRNA and up to five different Cas proteins. Identified invader DNA sequences are progressively unwound and cleaved by the action of the recruited nuclease and helicase Cas3 (*11, 12*).

CREDIT: B. STRAUCH/*SCIENCE*

Laboratory of Microbiology, Wageningen University, 6703 HB Wageningen, Netherlands.
E-mail: stan.brouns@wur.nl

is truly the Swiss army knife of CRISPR immunity.

REFERENCES AND NOTES

1. M. Jinek *et al.*, *Science* **337**, 816 (2012); 10.1126/science.1225829.
2. D. Bhaya, M. Davison, R. Barrangou, *Annu. Rev. Genet.* **45**, 273 (2011).
3. R. Barrangou *et al.*, *Science* **315**, 1709 (2007).
4. J. E. Garneau *et al.*, *Nature* **468**, 67 (2010).
5. R. Sapranauskas *et al.*, *Nucleic Acids Res.* **39**, 9275 (2011).
6. E. Deltcheva *et al.*, *Nature* **471**, 602 (2011).
7. G. Storz, J. Vogel, K. M. Wassarman, *Mol. Cell* **43**, 880 011).
8. H. Deveau *et al.*, *J. Bacteriol.* **190**, 1390 (2008).
9. E. C. Friedberg, *Nature* **421**, 436 (2003).
10. M. Baker, Nat. *Methods* **9,** 23 (2012).
11. B. Wiedenheft, S. H. Sternberg, J. A. Doudna, *Nature* **482**, 331 (2012).
12. E. R. Westra *et al.*, *Mol. Cell* **46**, 595 (2012).

10.1126/science.1227253

CRISPR/Cas, the immune system of bacteria and archaea

Philippe Horvath[1]* and Rodolphe Barrangou[2]*

Microbes rely on diverse defense mechanisms that allow them to withstand viral predation and exposure to invading nucleic acid. In many bacteria and most archaea, clustered regularly interspaced short palindromic repeats (CRISPR) form peculiar genetic loci, which provide acquired immunity against viruses and plasmids by targeting nucleic acid in a sequence-specific manner. These hypervariable loci take up genetic material from invasive elements and build up inheritable DNA-encoded immunity over time. Conversely, viruses have devised mutational escape strategies that allow them to circumvent the CRISPR/Cas system, albeit at a cost. CRISPR features may be exploited for typing purposes, epidemiological studies, host-virus ecological surveys, building specific immunity against undesirable genetic elements, and enhancing viral resistance in domesticated microbes.

Microbes have devised various strategies that allow them to survive exposure to foreign genetic elements. Although outpopulated and preyed upon by abundant and ubiquitous viruses, microbes routinely survive, persist, and occasionally thrive in hostile and competitive environments. The constant exposure to exogenous DNA via transduction, conjugation, and transformation have forced microbes to establish an array of defense mechanisms that allow the cell to recognize and distinguish incoming "foreign" DNA, from "self" DNA and to survive exposure to invasive elements. These systems maintain genetic integrity, yet occasionally allow exogenous DNA uptake and conservation of genetic material advantageous for adaptation to the environment. Certain strategies, such as prevention of adsorption, blocking of injection, and abortive infection, are effective against viruses; other defense systems specifically target invading nucleic acid, such as the restriction-modification system (R-M) and the use of sugar-nonspecific nucleases. Recently, an adaptive microbial immune system, clustered regularly interspaced short palindromic repeats (CRISPR) has been identified that provides acquired immunity against viruses and plasmids.

CRISPR represents a family of DNA repeats found in most archaeal (~90%) and bacterial (~40%) genomes (*1–3*). Although the initial discovery of a CRISPR structure was made fortuitously in *Escherichia coli* in 1987, the acronym was coined in 2002, after similar structures were observed in genomes of various bacteria and archaea (*1*). CRISPR loci typically consist of several noncontiguous direct repeats separated by stretches of variable sequences called spacers (which mostly correspond to segments of captured viral and plasmid sequences) and are often adjacent to *cas* genes (CRISPR-associated) (Fig. 1). *cas* genes encode a large and heterogeneous family of proteins that carry functional domains typical of nucleases, helicases, polymerases, and polynucleotide-binding proteins (*4*). CRISPR, in combination with Cas proteins, forms the CRISPR/Cas systems. Six "core" *cas* genes have been identified, including the universal markers of CRISPR/Cas systems *cas1* (COG1518) and *cas2* (COG1343, COG3512, occasionally in a fused form with other *cas* genes). Besides the *cas1* to *cas6* core genes, subtype-specific genes and genes encoding "repeat-associated mysterious proteins" (RAMP) have been identified and grouped into subtypes functionally paired with particular CRISPR repeat sequences (*4–8*). The size of CRISPR repeats and spacers varies between 23 to 47 base pairs (bp) and 21 to 72 bp, respectively. Generally, CRISPR repeat sequences are highly conserved within a given CRISPR locus, but a large assortment of repeat sequences has been shown across microbial species (*1, 9*). Most repeat sequences are partially palindromic, having the potential to form stable, highly conserved secondary structures (*7*). The number of repeat-spacer units is documented to reach 375 (*Chloroflexus* sp. Y-400-fl), but most loci commonly contain fewer than 50 units, as exemplified in lactic acid bacteria genomes (*8*). Microbes may contain more than one CRISPR locus; up to 18 such loci have been identified in *Methanocaldococcus jannaschii*, totaling more than 1% of the genome (*10*). CRISPRs are typically located on the chromosome, although some have been identified on plasmids (*11–13*).

The CRISPR loci have highly diverse and hypervariable spacer sequences, even between closely related strains (*14–16*), which were initially exploited for typing purposes. A variety of putative roles for CRISPR sequences was originally suggested, including chromosomal rearrangement, modulation of expression of neighboring genes, target for DNA binding proteins, replicon partitioning, and DNA repair (*5*). In 2005, three independent in silico studies reported homology between spacer sequences and extrachromosomal elements, such as viruses and plasmids (*11, 14, 15*). This led to the hypothesis that CRISPR may provide adaptive immunity against foreign genetic elements (*6*).

A vast spectrum of immunity

In 2007, it was shown in *Streptococcus thermophilus* that during natural generation of phage-resistant variants, bacteria commonly alter their CRISPR loci by polarized (i.e., at the leader end) incorporation of CRISPR repeat-spacer units (Fig. 1) (*17, 18*), consistent with observed spacer hypervariability at the leader end of CRISPR loci in various strains (*14, 16*). The integrated sequences were identical to those of the phages used in the challenge, which suggested that they originate from viral nucleic acid. To determine whether CRISPR impacts phage resistance, spacer content was altered via genetic engineering, which showed that spacer addition can provide novel phage resistance, whereas spacer deletion could result in loss of phage resistance (*17*). These findings were confirmed in *Streptococcus mutans*, where phage-resistant mutants acquired novel CRISPR spacers with sequences matching the phage genome, in vitro and in vivo (*19*). Although the ubiquitous and predatory nature of phages may explain the overwhelming representation of phage sequences in CRISPR loci, CRISPR spacers can also interfere with both plasmid conjugation and transformation, as shown in *Staphylococcus epidermidis* (*20*). Furthermore, several metagenomic studies investigating host-virus populations dynamics showed that CRISPR loci evolve in response to viral predation and that CRISPR spacer content and sequential order provide insights both historically and geographically (*21–24*).

The ability to provide defense against invading genetic elements seems to render CRISPR/Cas systems particularly desirable in hostile environments and may explain their propensity to be transferred horizontally between sometimes distant organisms (*12*). There is extensive evidence that defense systems such as CRISPR have undergone horizontal transfer between genomes, notably differences observed in codon bias, GC content variability, their presence on mobile genetic elements, the presence of neighboring insertion sequence elements, and their variable presence and location in closely related genomes. This is in agreement with the lack of congruence between the phylogenetic relation of various CRISPR elements and that of the organisms in which they are found (*8, 12*). This horizontal gene transfer may be mediated by plasmids, megaplasmids, and even prophages, all of which are documented to carry CRISPR loci (*2*).

Given the variety of defense systems in microbes and their role in controlling the presence of plasmids, prophages, transposons, and, perhaps, chromosomal sequences, studies should investigate whether CRISPR/Cas systems preferentially target certain elements and could determine whether they are symbiotic or mutually exclusive with other defense systems.

Idiosyncrasies of the CRISPR/Cas mechanism of action

The mechanism by which CRISPR provides resistance against foreign genetic elements is not fully characterized (Fig. 2). Even so, the functional link between Cas and CRISPR repeats has been inferred from the congruence observed between their sequence patterns. *cas* genes provide CRISPR-encoded immunity, because inactivating the CRISPR1-associated *cas7* gene (Fig. 1) impairs the ability of the host to integrate novel CRISPR spacers after phage exposure (*17*), which suggests that it is necessary for recognizing foreign nucleic acid and/or integrating the novel repeat-spacer unit. Cas1 appears to be a double-stranded DNA (dsDNA) endonuclease involved in the immunization process (*25*). It has also been proposed that Cas2 may act as a sequence-specific endoribonuclease that cleaves uracil-rich single-stranded RNAs (ssRNAs) (*26*). The mechanistic steps involved in invasive element recognition, novel repeat manufacturing, and spacer selection and integration into the CRISPR locus remain uncharacterized.

Although some Cas proteins are involved in the acquisition of novel spacers, others provide CRISPR-encoded phage resistance and interfere with invasive genetic elements. Mechanistically, although defense is spacer-encoded, the information that lies within the CRISPR repeat-spacer array becomes available to the Cas machinery through transcription. The CRISPR leader, defined as a low-complexity, A/T-rich, noncoding sequence, located immediately upstream of the first repeat, likely acts as a promoter for the transcription of the repeat-spacer array into a CRISPR transcript, the pre-crRNA (*13, 27*). The full-length pre-crRNA is subsequently processed into specific small RNA molecules that correspond to a spacer flanked by two partial repeats (*27–29*). In *E. coli*, processing is achieved by a multimeric complex of Cas proteins named Cascade (CRISPR-associated complex for antiviral defense), which specifically cleaves the pre-crRNA transcript within the repeat sequence to generate small CRISPR RNAs, crRNAs (*28*). Similarly, in *Pyrococcus*, Cas6 is an endoribonuclease that cleaves the pre-crRNA transcript into crRNA units that include a partial [8-nucleotide (nt)] repeat sequence at the 5′ end, as part of the Cas-crRNA complex (*27, 29, 30*). The crRNAs seem to specifically guide the Cas interference machinery toward foreign nucleic acid molecules that match its sequence, which leads ultimately to degradation of the invading element (*30*). The involvement of *cas* genes in CRISPR defense was originally demonstrated when inactivating the CRISPR1-associated *csn1*-like gene (Fig. 1) resulted in loss of phage resistance despite the presence of matching spacers (*17*).

The observation that CRISPR spacers match both sense and antisense viral DNA led to the hypothesis that some CRISPR/Cas systems may target dsDNA, and this was confirmed by disruption of target DNA with an intron (the excision of which restores the native mRNA) on a plasmid that allows conjugation despite the presence of a matching CRISPR spacer (*20*). Conversely, the *Pyrococcus* CRISPR effector complex, a

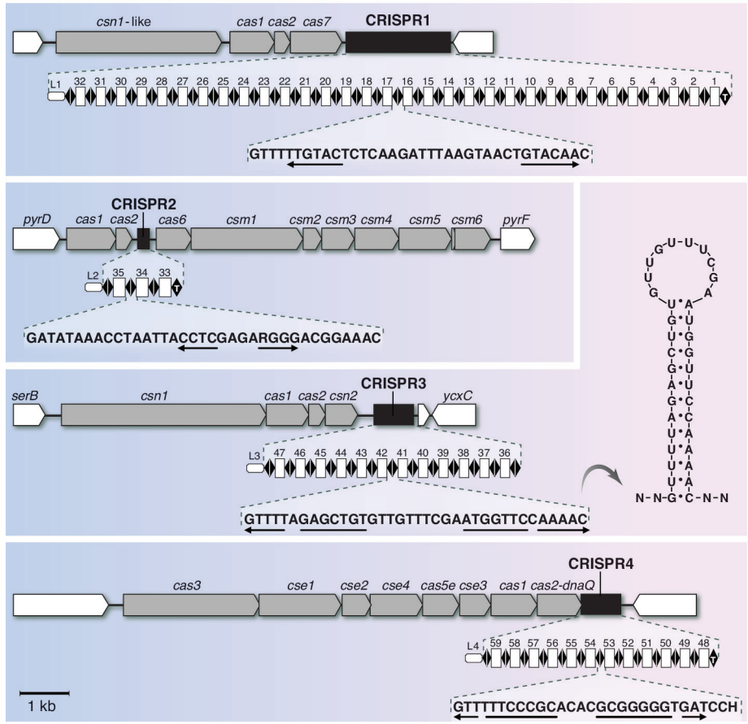


Fig. 1. Overview of the four CRISPR/*cas* systems present in *Streptococcus thermophilus* DGCC7710. For each system, gene organization is depicted on the top, with *cas* genes in gray, and the repeat-spacer array in black. Below the gene scheme, the repeat and spacer (captured phage or plasmid nucleic acid) content is detailed as black diamonds (T, terminal repeat) and white rectangles, respectively. Bottom line, consensus repeat sequence. L1 to L4, leader sequences. The predicted secondary structure of the CRISPR3 repeat is shown on the right. *S. thermophilus* CRISPR2, CRISPR3, and CRISPR4 systems are homologous to the CRISPR systems of *Staphylococcus epidermidis* (*20*), *Streptococcus mutans* (*19*), and *E. coli* (*28*), respectively.

[1]Danisco France SAS, BP10, F-86220 Dangé-Saint-Romain, France.
[2]Danisco USA, Inc., 3329 Agriculture Drive, Madison, WI 53716, USA.
*Corresponding author. E-mail: philippe.horvath@danisco.com (P.H.), rodolphe.barrangou@danisco.com (R.B.)

ribonucleoprotein complex that consists of crRNA and Cas proteins, targets invader RNA by complementary-dependent cleavage, in vitro (*30*). Given the large diversity of CRISPR/Cas systems in bacteria and archaea (*4, 6*), it is likely that both DNA and RNA may be targets. More information is needed to establish and understand what the functional differences are among distinct CRISPR/Cas systems.

The initial hypothesis that CRISPR may mediate microbial immunity via RNA interference (RNAi) (*6*) is misguided. RNAi allows eukaryotic organisms exposed to foreign genetic material to silence the invading nucleic acid sequence before or after it integrates into the host chromosome, and/or to subvert cellular processes through a small interfering RNA guide (*31*). A key difference between RNAi and CRISPR-encoded immunity lies in the enzymatic machinery involved. Although both are mediated by a guide RNA in an inhibitory ribonucleoprotein complex, only Dicer, Slicer, and the RNA-induced silencing complex (RISC) may have analogous counterparts (*6, 30*). Mechanistically, although the short RNA duplexes at the core of RNAi are typically 21 to 28 nt in length (*32*), crRNAs are larger, because they contain a CRISPR spacer (23 to 47 nt) flanked by partial repeats. Also, RNA-dependent transcription generating dsRNA and using the cleaved target RNA seen in RNAi have not been characterized in the CRISPR/Cas systems. In other ways, the sequence-specific and adaptive CRISPR/Cas systems share similarities with the vertebrate adaptive immune system, although CRISPR spacers are DNA-encoded and can be inherited by the progeny.

Circumventing CRISPR-based immunity

Even though CRISPR can provide high levels of phage resistance, a relatively small proportion of viruses retain the ability to infect the "immunized" host. These viral particles have specifically mutated the proto-spacer (sequence within the invading nucleic acid that matches a CRISPR spacer), with a single point mutation that allows the viruses to overcome immunity, which indicates that the selective pressure imposed by CRISPR can rapidly drive mutation patterns in viruses (*17, 18, 23*). Analysis of phage sequences adjacent to proto-spacers revealed the presence of conserved sequences, called CRISPR motifs (*13, 16, 18, 19, 33, 34*), or proto-spacer adjacent motifs (PAMs) (*35*). Phages may also circumvent the CRISPR/Cas system by mutating the CRISPR motif (*18*), which indicates that it is involved in CRISPR-encoded immunity. Additionally, CRISPR motif mutation can result in loss of phage resistance despite the presence of a matching CRISPR spacer (*34*). The absence of this motif in the CRISPR locus likely allows the system to act on the invading target DNA specifically and precludes an "autoimmune" response on the host chromosome (Fig. 2). Such a motif may not be necessary in CRISPR/Cas systems targeting RNA. Although proto-spacers seem to be randomly located on phage genomes, a given CRISPR spacer may be acquired independently by different lineages. It is thus tempting to speculate that CRISPR motifs also play a key role in the selection of spacers.

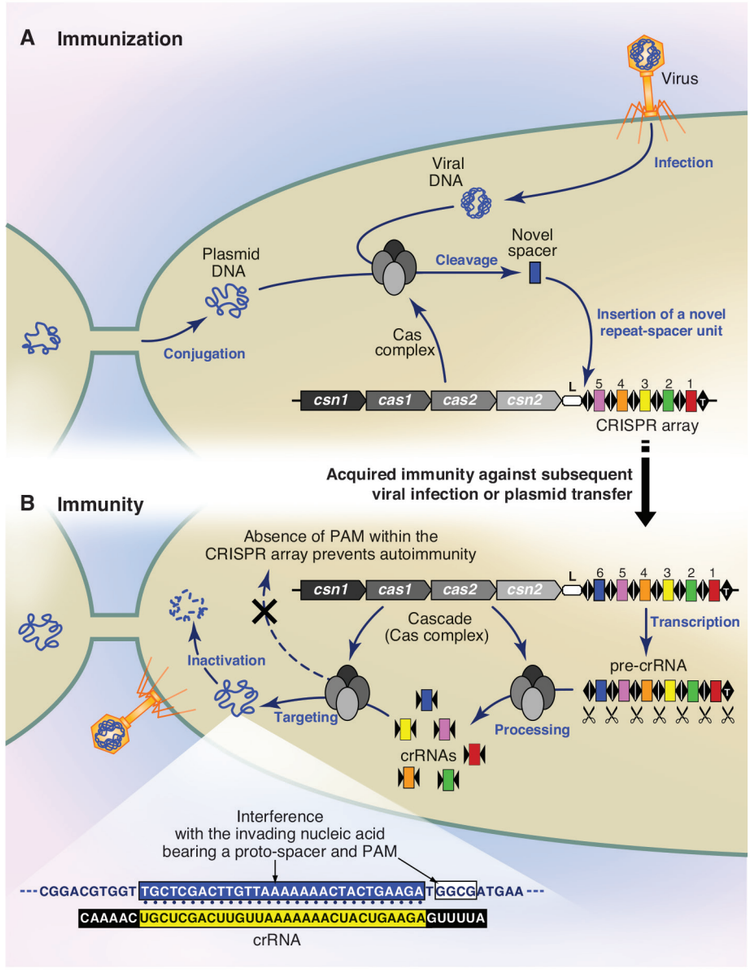


Fig. 2. Overview of the CRISPR/Cas mechanism of action. (**A**) Immunization process: After insertion of exogenous DNA from viruses or plasmids, a Cas complex recognizes foreign DNA and integrates a novel repeat-spacer unit at the leader end of the CRISPR locus. (**B**) Immunity process: The CRISPR repeat-spacer array is transcribed into a pre-crRNA that is processed into mature crRNAs, which are subsequently used as a guide by a Cas complex to interfere with the corresponding invading nucleic acid. Repeats are represented as diamonds, spacers as rectangles, and the CRISPR leader is labeled L.

These mutations may have an impact on the amino acid sequence, as either nonsynonymous mutations or premature stop codons that truncate the viral protein (*18*). In addition to mutations, phages may also circumvent CRISPR-encoded immunity via deletion of the target sequence (*18, 21*). This perhaps indicates a strong cost associated with circumventing the CRISPR/Cas systems. Alternative strategies that allow viruses to escape CRISPR, such as suppressors that could interfere with crRNAs biogenesis or Cas machinery remain uncovered. Defense tactics employed by viruses to circumvent the CRISPR/Cas systems are yet another critical difference between RNAi and CRISPR: Eukaryotic viruses may express inhibitors such as dsRNA-binding proteins that interfere with the RNA silencing machinery (*32*), which are yet to be identified in response to CRISPR, whereas microbial viruses specifically mutate or recombine (*21*) the sequence corresponding to the CRISPR spacer or that of the PAM.

The impact of CRISPR on phage genomes is illustrated by extensive genome recombination events observed in environmental phage populations in response to CRISPR (*21*). This contrasts with the fact that acquisition of novel CRISPR spacers does not seem to have a fitness cost for the host, apart from maintaining the CRISPR/Cas system as active.

Although it seems intuitive that CRISPR loci should not be able to expand indefinitely (*21, 36*), little is known about the parameters that define the optimal and maximum size of a CRISPR locus. Also, the fitness cost of CRISPR expansion in the host should be compared with that of CRISPR evasion in the virus populations, so as to determine whether prey or predators incur the higher evolutionary cost of this genetic warfare.

Although CRISPR loci primarily evolve via polarized addition of novel spacers at the leader end of the locus after phage exposure, internal spacer deletions have also been reported, likely occurring via homologous recombination between CRISPR repeats (*1, 16, 18*). Perhaps this allows the host to limit the expansion of the CRISPR locus so that the relative size of the locus does not increase to a detrimental level. The propensity of spacers located at the trailer end (opposite to the leader end) to be deleted preferentially would mitigate the loss of fitness associated with the deletion, because ancestral spacers would arguably provide resistance against viruses that were historically, but are not currently, present in the environment. The combination of locus expansion via spacer acquisition and contraction via spacer loss, in the context of rapid evolution in space and time because of viral predation, which generate a high level of spacer polymorphism, suggests that CRISPR loci undergo dynamic and rapid turnover on evolutionary time scales (*16, 21, 36*). Indeed, in microbes with an active system, CRISPR loci have been shown to be the most hypervariable genomic regions (*21*).

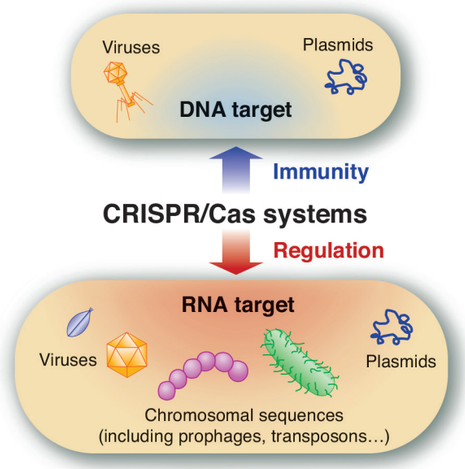


Fig. 3. CRISPR interference. The CRISPR/Cas systems may target either DNA or RNA to interfere with viruses, plasmids, prophages, or other chromosomally encoded sequences.

Applications and future directions

A priori, the concurrent presence of distinct defense systems against foreign genetic elements in bacteria and archaea seems inefficient and redundant, although it might reflect functional preferences and increased fitness. Because all defense mechanisms have their advantages and caveats, the accumulation and combination of different systems would increase the selective pressure on invading elements and, consequently, could increase the chances of host survival by using multiple hurdles.

Because CRISPR spacers correspond to prior episodes of phage and plasmid exposure, they provide a historical and geographical—although limited—perspective as to the origin and paths of a particular strain, which may be used for ecological and epidemiological studies. Many intrinsic aspects of CRISPR-based immunity have provided avenues for industrial applications, including exploiting hypervariability for typing purposes, driving viral evolution, predicting and modulating virus resistance in domesticated microbes, and performing natural genetic tagging of proprietary strains. The inheritable nature of the CRISPR spacer content provides potential for perennial use of industrial microbes. Alternatively, the ability of CRISPR/Cas systems to impede the transfer of particular nucleic acid sequences (such as phage or plasmid DNA) into a host might be exploited via genetic engineering to specifically preclude the dissemination of undesirable genetic elements, such as antibiotic-resistance markers and genes harmful to humans and other living organisms. It may also be designed to limit the intracellular spread of mobile genetic elements such as insertion sequences and transposons. In addition to providing immunity, CRISPR/Cas systems that target RNA have the potential to affect the transcript stability of chromosomal elements (Fig. 3).

Although significant progress has been made in the last few years, many mechanistic aspects remain uncovered, notably vis-à-vis the immunization process (key elements involved in spacer selection and integration between repeats and/or possible involvement of degenerate infectious particles in building immunity) and the interference mechanism (other cellular components involved). Also, more knowledge is desirable regarding the elements necessary to have functional CRISPR/Cas systems and the basis for the absence of CRISPR in 60% of bacteria.

REFERENCES AND NOTES

1. R. Jansen, J. D. Embden, W. Gaastra, L. M. Schouls, *Mol. Microbiol.* **43**, 1565 (2002).
2. R. Sorek, V. Kunin, P. Hugenholtz, *Nat. Rev. Microbiol.* **6**, 181 (2008).
3. J. van der Oost, M. M. Jore, E. R. Westra, M. Lundgren, S. J. Brouns, Trends Biochem. *Sci.* **34**, 401 (2009).
4. D. H. Haft, J. Selengut, E. F. Mongodin, K. E. Nelson, *PLOS Comput. Biol.* **1**, e60 (2005).
5. K. S. Makarova, L. Aravind, N. V. Grishin, I. B. Rogozin, E. V. Koonin, *Nucleic Acids Res.* **30**, 482 (2002).
6. K. S. Makarova, N. V. Grishin, S. A. Shabalina, Y. I. Wolf, E. V. Koonin, *Biol. Direct* **1**, 7 (2006).
7. V. Kunin, R. Sorek, P. Hugenholtz, *Genome Biol.* **8**, R61 (2007).
8. P. Horvath *et al.*, *Int. J. Food Microbiol.* **131**, 62 (2009).
9. I. Grissa, G. Vergnaud, C. Pourcel, *BMC Bioinformatics* **8,** 172 (2007).
10. R. K. Lillestøl, P. Redder, R. A. Garrett, K. Brügger, *Archaea* **2**, 59 (2006).
11. F. J. Mojica, C. Díez-Villaseñor, J. García-Martínez, E. Soria, *J. Mol. Evol.* **60**, 174 (2005).
12. J. S. Godde, A. Bickerton, *J. Mol. Evol.* **62**, 718 (2006).
13. R. K. Lillestøl *et al.*, *Mol. Microbiol.* **72**, 259 (2009).
14. C. Pourcel, G. Salvignol, G. Vergnaud, *Microbiology* **151**, 653 (2005).
15. A. Bolotin, B. Quinquis, A. Sorokin, S. D. Ehrlich, *Microbiology* **151**, 2551 (2005).
16. P. Horvath *et al.*, *J. Bacteriol.* **190**, 1401 (2008).
17. R. Barrangou *et al.*, *Science* **315**, 1709 (2007).
18. H. Deveau *et al.*, *J. Bacteriol.* **190**, 1390 (2008).
19. J. R. van der Ploeg, *Microbiology* **155**, 966 (2009).
20. L. A. Marraffini, E. J. Sontheimer, *Science* **322**, 1843 (2008).
21. A. F. Andersson, J. F. Banfield, *Science* **320**, 1047 (2008).
22. V. Kunin *et al.*, *Genome Res.* **18**, 293 (2008).
23. J. F. Heidelberg, W. C. Nelson, T. Schoenfeld, D. Bhaya, N. Ahmed, *PLOS ONE* **4**, e4169 (2009).
24. N. L. Held, R. J. Whitaker, *Environ. Microbiol.* **11**, 457 (2009).
25. B. Wiedenheft *et al.*, *Structure* **17**, 904 (2009).
26. N. Beloglazova *et al.*, *J. Biol. Chem.* **283**, 20361 (2008).
27. C. Hale, K. Kleppe, R. M. Terns, M. P. Terns, *RNA* **14**, 2572 (2008).
28. S. J. Brouns *et al.*, *Science* **321**, 960 (2008).
29. J. Carte, R. Wang, H. Li, R. M. Terns, M. P. Terns, *Genes Dev.* **22**, 3489 (2008).
30. C. R. Hale *et al.*, *Cell* **139**, 945 (2009).
31. C. C. Mello, D. Conte Jr., *Nature* **431**, 338 (2004).
32. G. Meister, T. Tuschl, *Nature* **431**, 343 (2004).
33. S. A. Shah, N. R. Hansen, R. A. Garrett, Biochem. *Soc. Trans.* **37**, 23 (2009).
34. E. Semenova, M. Nagornykh, M. Pyatnitskiy, I. I. Artamonova, K. Severinov, *FEMS Microbiol. Lett.* **296**, 110 (2009).
35. F. J. Mojica, C. Díez-Villaseñor, J. García-Martínez, C. Almendros, *Microbiology* **155**, 733 (2009).
36. G. W. Tyson, J. F. Banfield, Environ. *Microbiol.* **10**, 200 (2008).

37. We thank our colleagues P. Boyaval, C. Fremaux, D. Romero, and E. Bech Hansen for their support and scientific contributions, and S. Moineau, V. Siksnys, and J. Banfield for their insights and expertise. This work was supported by Danisco A/S. P.H. and R.B. have submitted patent applications relating to various uses of CRISPR.

10.1126/science.1179555

ABSTRACT ONLY

A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity

Martin Jinek,[1,2*] Krzysztof Chylinski,[3,4*] Ines Fonfara,[4] Michael Hauer,[2†] Jennifer A. Doudna,[1,2,5,6‡] Emmanuelle Charpentier[4‡]

Clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems provide bacteria and archaea with adaptive immunity against viruses and plasmids by using CRISPR RNAs (crRNAs) to guide the silencing of invading nucleic acids. We show here that in a subset of these systems, the mature crRNA that is base-paired to trans-activating crRNA (tracrRNA) forms a two-RNA structure that directs the CRISPR-associated protein Cas9 to introduce double-stranded (ds) breaks in target DNA. At sites complementary to the crRNA-guide sequence, the Cas9 HNH nuclease domain cleaves the complementary strand, whereas the Cas9 RuvC-like domain cleaves the noncomplementary strand. The dual-tracrRNA:crRNA, when engineered as a single RNA chimera, also directs sequence-specific Cas9 dsDNA cleavage. Our study reveals a family of endonucleases that use dual-RNAs for site-specific DNA cleavage and highlights the potential to exploit the system for RNA-programmable genome editing.

[1]Howard Hughes Medical Institute (HHMI), University of California, Berkeley, CA 94720, USA.
[2]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA.
[3]Max F. Perutz Laboratories (MFPL), University of Vienna, A-1030 Vienna, Austria.
[4]The Laboratory for Molecular Infection Medicine Sweden, Umeå Centre for Microbial Research, Department of Molecular Biology, Umeå University, S-90187 Umeå, Sweden.
[5]Department of Chemistry, University of California, Berkeley, CA 94720, USA.
[6]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.
*These authors contributed equally to this work.
†Present address: Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland.
‡Corresponding author.
E-mail: doudna@berkeley.edu (J.A.D.); emmanuelle.charpentier@mims.umu.se (E.C.)

Read full article: scim.ag/1piiXv7

Originally published 17 August 2012

Real-World Gene Editing Issues: Improving Knock-In Efficiency

Authors: Tom Henley, Ramu Mangena, Rohan Sivapalan, Eric Rhodes

Genome editing has gotten a real boost from the recent interest in CRISPR technology. Unfortunately while many experimenters are having great success using CRISPR to knock-out genes, the ability to use CRISPR for knock-ins has proven somewhat more challenging.

The difficulties associated with efficiently generating knock-ins can be attributed to several issues, but two are key: the need to position the cut site close to the region of the intended change and the need for effective delivery of a donor DNA template. These constraints generally don't exist when doing targeted knock-outs but present several 'real-world' problems when attempting to generate knock-ins.

It has been demonstrated that the efficiency of incorporation of a targeted alteration is highly dependent on the distance from the nuclease-induced cut site to the site of the desired genomic alteration[1]. The closer these two are to each other, the higher the efficiency, so it would make sense to use a guide RNA (gRNA) that lies as close to the site of the desired alteration as possible. However, the best gRNA to drive incorporation may not be the most specific, leading to off-target effects, and so a difficult choice needs to be made as to the right balance to strike between these two factors.

It has also been demonstrated that not all gRNAs are equal in activity. Therefore, the lack of sufficient activity by a given gRNA might also require one to use a gRNA which lies somewhat further away from the target site. It is even possible, in some cases, that there simply is no suitable gRNA available that lies close enough (within 40bp) to the target site.

To address the donor side of the equation, researchers have mostly been using either single-stranded oligos or double-stranded plasmids to provide a donor template. Because of Horizon's long history of using rAAV (recombinant adeno-associated virus) to edit and create hundreds of isogenic lines, we were interested to see whether the combination of rAAV and CRISPR might offer advantages not necessarily realized when using an oligo or plasmid alone.

It has been known for some time that rAAV is able to drive homologous recombination levels that are up to 1,000-fold higher than those seen when using a simple plasmid. Despite this vast improvement over plasmids, the levels of recombination seen are still relatively modest, making rAAV optimal for single allele alterations, but less so when looking to affect multiple alleles simultaneously. The efficiency of using rAAV is boosted significantly however, when a double-strand break is introduced in the vicinity of the homology region[2]. In addition, rAAV has an extremely wide tropism making it very effective for delivery of a single-stranded piece of long DNA (approx. 4.8kb) directly to the nucleus of many cell types, particularly those which may be difficult to transfect.

To test the potential for a combination approach, we set up an experiment to look at more "real-world" situations where the gRNA cut wasn't necessarily located directly over the desired change and used a common cell line (HCT116) that is not particularly easy to transfect. We used CRISPR to introduce double strands breaks at various points in the genome of a line carrying a disabled copy of the GFP gene on one chromosome. The figure below shows the organization of the artificial GFP gene split into three exons, with the third exon carrying a mutation which prematurely terminates the GFP protein rendering it non-functional. Conversion of the adenosine (A) residue to a cysteine (C) restores GFP functionality. The location and relative distance of five different gRNA targets sites from the point mutation are shown below the expanded Exon 3 region.

Figure 1: *An inactive GFP gene, split into exons, driven by the CMV promoter is integrated as a single copy into the genome. gRNA target sites are shown flanking the mutation which renders the GFP non-functional, with numbers representing the distance of each PAM site from the mutation.*

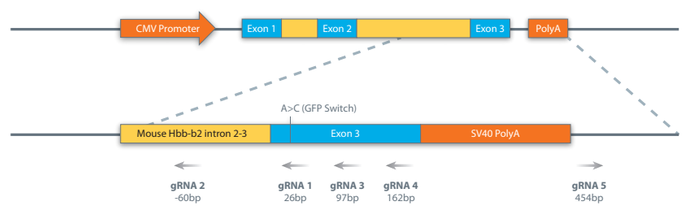


Figure 2: *Plasmid and rAAV donors were compared for their ability to knock-in the activating mutation and restore GFP function through homology-directed repair. For each gRNA tested, rAAV donors provide higher frequencies of correct gene targeting than equivalent plasmids.*

We used both a plasmid and rAAV derived from that same plasmid as a donor and measured the ability of these donors to stimulate a genetic modification at the range of distances from the induced cut site.

Remarkably the rAAV donor performs markedly better than the corresponding plasmid with an average 10-fold improvement at most distances. We are currently working to extend this research and recent results comparing rAAV with oligo and plasmid donors at different concentrations have supported this trend.

We acknowledge that the data is representative of only one cell line which does not transfect particularly well, but we feel it does fairly represent a real-world situation which often arises as experiments are designed in an ever widening variety of cellular backgrounds. We and other groups continue to look for technology combinations and other approaches that can be used to boost gene targeting efficiencies and lower the barriers to make even complex gene editing more routine.

References

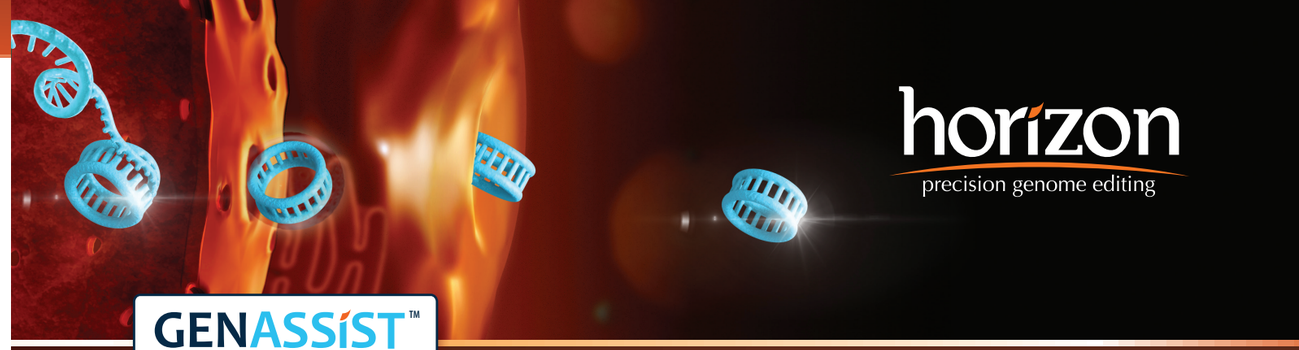
1 L. Yang *et al.*, *Nucleic Acids Res*. 41, 9049 (2013).

2 M. H. Porteus *et al.*, Molec. Cell. Biol. 23, 3558 (2003).

Contact us to learn more about Horizon's research:

Website: www.horizondiscovery.com
Email: info@horizondiscovery.com
Phone: UK +44 (0)1223 655 580

horizon
precision genome editing



horizon
precision genome editing

GENASSIST™
PRECISION GENOME EDITING TOOLS

GENASSIST™ | CRISPR & rAAV Genome Editing Tools

A range of gene-editing kits and reagents to enable easier, robust implementation of CRISPR and rAAV gene-editing experiments. The GENASSIST™ product and service solution includes:

- Custom guide RNA manufacturing **and Validation**
- Custom **Donor Design** and manufacturing
- Both **CRISPR and rAAV** technologies for the right solution to your challenge
- **Adenovirus and rAAV** packaging for hard to transfect cells
- Consultation with our **Gene-Editing Expert Scientists**
- Over 550 genetically defined **Isogenic Cell Lines** available for further genomic modification
- **And** an extensive plasmid library!

Generate the cell line you want in your lab:

- Point Mutations
- Gene Knockouts
- Deletions
- Insertions
- Translocations
- Amplifications

Or choose any of our over 550 off the shelf X-MAN™ Isogenic Cell Line pairs to start driving your research right away!

Visit www.horizondiscovery.com/GENASSIST to learn more about GENASSIST™ and Horizon's many other products and services

horizon
precision genome editing

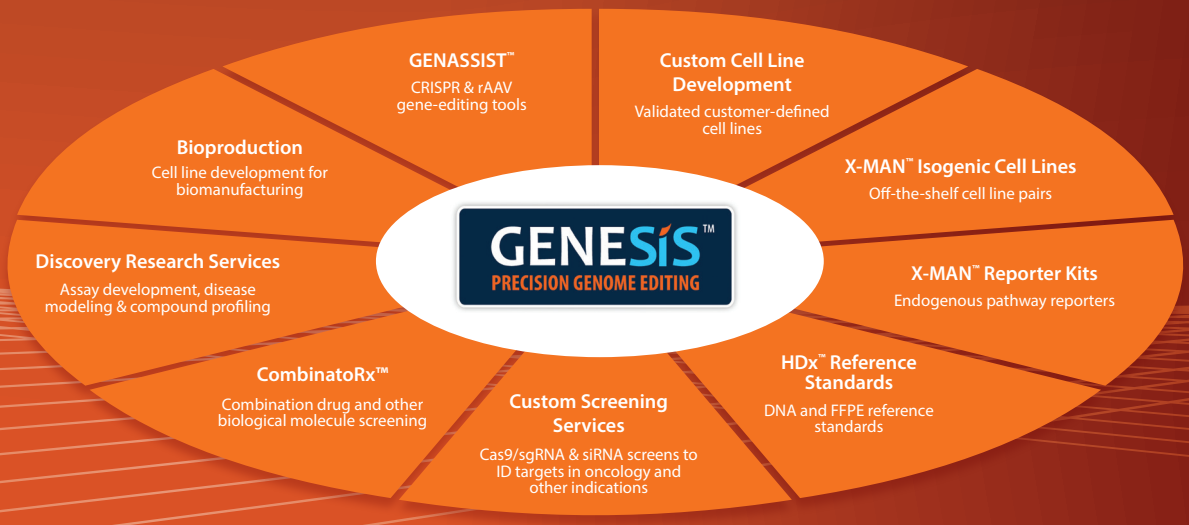
Translating Genomes | Personalizing Medicine

Horizon Discovery combines long scientific heritage in translational research with GENESIS™, a precision gene-editing platform incorporating rAAV, CRISPR and ZFN technologies. Horizon supplies genetically-defined cell lines, gene-editing tools and services, custom cell line generation, molecular reference standards, and contract research services to almost 1,000 academic, clinical and biopharmaceutical organizations.

Horizon's precision gene-editing platform

As the only translational genomics company to have access to all three gene-editing technologies, Horizon is in a unique position to be able to generate virtually any genomic modification upon customer request.

Horizon's Capabilities:



For more information or to learn more about Horizon's products & services:

Visit www.horizondiscovery.com